



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
"Харківський політехнічний інститут"

О. Б. Білоцерківський, Н. В. Ширяєва, О. О. Замула

ЕКОНОМІКО-МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Текст лекцій

Харків 2010

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
"ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"

О. Б. Білоцерківський, Н. В. Ширяєва, О. О. Замула

ЕКОНОМІКО-МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Текст лекцій

Затверджено
редакційно-видавничою
радою університету,
протокол № 2 від 18.09.2009

Харків НТУ „ХПІ” 2010

ББК 22.213

Б 78

УДК 534.1

Рецензенти: *В.А. Смоляк*, канд. екон. наук, доц. ХНЕУ;

О.М. Гаврись, канд. екон. наук, проф. НТУ "ХПІ"

Б 78 Білоцерківський О. Б. Економіко-математичне моделювання : Текст лекцій / О. Б. Білоцерківський, Н. В. Ширяєва, О. О. Замула. – Х.: НТУ "ХПІ", 2010. – 108 с.

ISBN

Текст лекцій містить основи економіко-математичного моделювання, включаючи задачі лінійного, цілочислового, нелінійного та динамічного програмування, елементи теорії ігор, парний та множинний регресійний аналізи, нелінійні рівняння, часові ряди та їх використання в економічних дослідженнях. Надано типові приклади з розв'язаннями за матеріалом, що вивчається.

Призначено для студентів спеціальностей 7.050206 "Менеджмент зовнішньоекономічної діяльності" та 6.030508 "Фінанси".

Іл. 10. Табл. 15. Бібліогр. 12 назв.

ISBN

ББК 22.213

© О. Б. Білоцерківський,
Н. В. Ширяєва, О. О. Замула,
2010 р.

ВСТУП

В умовах України як самостійної держави зростає роль економіко-математичних методів як одного із способів розвитку динамічно розвинутої та стійкої економіки з науково обґрунтованими шляхами розвитку та прогнозами на майбутнє в перехідному етапі до ринку.

Економіко-математичні методи – це узагальнена назва комплексу економіко-математичних підходів, об'єднаних для вивчення економіки та призначених для побудови, реалізації і дослідження економічних моделей.

Математичне моделювання в економіці та менеджменті – це використання математичного моделювання в розв'язанні господарських завдань і обґрунтуванні прийнятих рішень щодо керування виробництвом.

У даному тексті лекцій викладаються теоретичні основи курсу «Економіко-математичне моделювання», включаючи задачі лінійного, цілочислового, нелінійного та динамічного програмування, елементи теорії ігор, парний та множинний регресійний аналізи, нелінійні рівняння, часові ряди та їх використання в економічних дослідженнях. Текст лекцій містить достатню кількість розв'язаних завдань.

Цей текст лекцій розрахований на студентів спеціальностей 7.080206 "Менеджмент зовнішньоекономічної діяльності" та 6.030508 "Фінанси", що вивчають курс "Економіко-математичне моделювання". Він також буде корисний для студентів інших економічних спеціальностей.

1. КОНЦЕПТУАЛЬНІ АСПЕКТИ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ ЕКОНОМІКИ

1.1. Предмет, мета, завдання та основні поняття економіко-математичного моделювання

Предметом економіко-математичного моделювання є методологія та інструментарій побудови і розв'язування детермінованих оптимізаційних задач.

Мета – формування системи знань з методології та інструментарію побудови і використання різних типів економіко-математичних моделей.

Завдання – вивчення основних принципів та інструментарію постановки задач, побудови економіко-математичних моделей, методів їх розв'язування та аналізу з метою використання в економіці.

Основні поняття економіко-математичного моделювання

Мета – це фундаментальне поняття, тому що економічна діяльність завжди цілеспрямована. Під метою розуміють бажаний результат, що повинний бути досягнутий.

Захід – сукупність дій, об'єднаних загальною метою. В дослідженні операцій (відгалуженні кібернетики) замість терміну «захід» використовується поняття «операція».

Альтернативи – можливі варіанти заходів, на підставі яких приймається рішення. Таких варіантів може бути декілька. Альтернативи можуть бути дискретними або безперервними. Кількість дискретних альтернатив скінченна: наприклад, замінити певний вид устаткування або ні (у даному випадку альтернативи дві). Альтернативи можуть вибиратися на безперервній множині: наприклад, коли замінити устаткування даного виду (через день, два, тиждень, місяць, рік і т.д.), тоді кількість альтернатив нескінченна і під рішенням розуміють вибір однієї альтернативи з безліч можливих.

Система (у перекладі з грецької – ціле, складене з частин) – це множина взаємозв'язаних елементів, які становлять певну єдність.

Елемент системи – частина системи, яка, виходячи з мети та функцій даної системи, є неподільною.

Складна система – це безліч різних структур і елементів цих структур.

Підсистема – частина системи, яка виділена з певною метою, може розглядатися як самостійна система.

Системний підхід – головний науковий принцип дослідження систем, згідно з яким необхідно враховувати взаємозв'язки між елементами системи, між системою та зовнішнім середовищем, між станом системи у даний час і майбутньому. Основне поняття в кібернетиці.

Модель – система, здатна замінити оригінал (тобто реальну систему) так, що її вивчення дає інформацію про оригінал. Модель може повністю або частково відтворювати структуру системи, що моделюється, та її функції.

Моделювання – процес побудови, реалізації та дослідження моделі, який здатний замінити реальну систему та дати інформацію про неї.

Математична модель – система математичних і логічних співвідношень, які описують структуру та функції реальної системи. Математична модель відрізняється за своєю природою від оригіналу. Дослідження властивостей оригіналу за допомогою математичної моделі зручніше, дешевше, забирає менш часу порівняно з фізичним моделюванням, яке використовується в техніці (тобто має ту ж природу, що і оригінал). Більш того, цілий ряд економічних систем неможливо зобразити за допомогою фізичних моделей.

Економіко-математична модель – це математичний опис економічного процесу чи явища з метою його дослідження та керування. Вона включає в себе систему рівнянь та нерівностей математичного опису економічних процесів і явищ, які складаються з набору змінних і параметрів. *Змінні величини* характеризують, наприклад, обсяг виробленої продукції, капітальних вкладень, перевезень тощо. Змінні поділяються на дві групи: *пояснювальні* (залежні), які є наперед заданими та незалежними, *пояснювані* (незалежні), які є результативними показниками. Змінні величини можуть бути двох груп: *зовнішні змінні* (*екзогенні*), коли вони визначаються поза даною моделлю та вважаються для моделі заданими, *внутрішні змінні* (*ендогенні*), які визначаються в результаті дослідження даної моделі. *Параметри* – це чисельні ознаки показників, такі, як норми витрат сировини, матеріалів, часу на виробництво тощо. В усіх випадках необхідно, щоб модель мала достатньо детальний опис об'єкта, який дозволяв би здійснювати вимір економічних величин та їх взаємозв'язок, щоб були виділені фактори, які впливають на досліджувані показники.

Оптимізаційна модель дозволяє з декількох альтернативних варіантів вибрати найкращий варіант за будь-якою ознакою.

Економіко-математичні методи – узагальнена назва комплексу економіко-математичних підходів, об'єднаних для вивчення економіки та призначених для побудови, реалізації і дослідження економічних моделей.

1.2. Історія розвитку економіко-математичних методів

XVIII ст. – початок використання математичних методів в економіці з опублікування роботи *"Економічні таблиці"* французьким економістом *Ф. Кене*, який вперше зробив спробу формалізації процесу суспільного відтворення. В подальшому наукове обґрунтування суспільного відтворення було здійснено *К. Марксом*.

XIX ст. – формується економетрія як наука з початку розробки статистичних методів у вигляді парної та множинної регресії, теорії кореляції, теорії помилок, вибірових методів (*Р. Гамільтон, К. Пірсон, Р. Фішер та ін.*).

У середині 30–40-х років XX ст. виникають лінійні методи оптимізації: *лінійне програмування*, скорочено ЛП (*Л.В. Канторович, Дж. Данциг*) та *теорія ігор* (*Дж. фон Нейман*).

Досвід використання лінійних моделей показав, що вони далеко не завжди можуть бути використані для опису економічних процесів і явищ. Тому починають розвиватися дослідження в інших напрямках *нелінійного програмування*: випуклого, геометричного, динамічного та ін.

У 1948 р. виникає нова наука – *кібернетика* (у перекладі з грецької – мистецтво керування), засновником якої став американський математик *Норберт Вінер*. Кібернетика – це наука про загальні закономірності процесів керування в різних системах: біологічних, економічних, технічних та ін. Одним з напрямків кібернетики, об'єктом якого постають економічні системи, є *економічна кібернетика*. Вона лежить в основі побудови ряду оптимізаційних моделей. Пізніше розвиваються такі прикладні напрямки економічної кібернетики, як *дослідження операцій* (пошук шляхів раціонального використання ресурсів для реалізації поставлених цілей), *теорія масового обслуговування* (яка розглядає різні явища в економіці – процеси обслуговування, тобто задовільнення будь-яких запитів, замовлень тощо).

У 50–60-х роках макроекономічні дослідження в економетрії проводять *Я. Тінберген, Р. Фріш*. Центром розвитку економетрії стала Комісія Коулса (США). Новий інструментарій економетрії одержала в результаті розробки моделей одночасних рівнянь (*Т. Хаавельмо, Т. Купманс, Г. Гейл та ін.*). Серед но-

вих економетричних систем, за якими розрахунки починають вестись з використанням ЕОМ, виникають такі макроекономічні моделі: *Брукінгська модель (США), Голландська модель, Уортонська модель (США)*, які використовуються для прогнозування та розробки економічної політики, для аналізу попиту та споживання.

У 60-х роках починається впровадження в практику планування СРСР нових методів, які одержали назву "*Сіткові методи планування та управління*" (СПУ). Вони лежать в основі сіткових моделей.

Набувають розвитку деякі розділи прикладної математики, які зв'язані з вирішенням оптимізаційних задач: *нелінійне математичне програмування, математична теорія оптимізаційних процесів*.

Відповідний внесок у розвиток економетрії роблять *вітчизняні вчені-економісти (Є.Є. Слуцький, Л.В. Канторович, В.С. Немчинов та ін.)*. Так, акад. В.С. Немчинову належить значна роль у реабілітації в СРСР існуючого погляду на економетрію як "буржуазну", "антимарксистську" та "шкідливу лже-науку" (1965 р.).

У 70–90-х роках економіко-математичне моделювання стало визнаним способом аналізу економічних проблем. У вітчизняній практиці у 70-х роках з'являються *автоматизовані системи управління (АСУ)*, призначені для оптимізації керування складними виробничими процесами та економічними системами.

У наш час набувають впровадження у вітчизняну практику *економетричні підходи з використанням програмних комплексів ПК*. В умовах України як самостійної держави зростає роль економіко-математичних методів як одного із способів розвитку динамічно розвинутої та стійкої економіки з науково обґрунтованими шляхами розвитку та прогнозами на майбутнє в перехідному етапі до ринку.

1.3. Сучасний стан економіко-математичного моделювання

У теперішній час сфера можливого використання економіко-математичних методів і моделей у плануванні та керуванні значна і з кожним роком вона розширюється, але галузь фактичного їх використання на практиці пов'язана з *труднощами*:

✓ **складність моделювання** економічних процесів і явищ з урахуванням виробничих відносин (поведінка людей, їх інтереси, індивідуально прийняті рішення);

✓ **необхідність "вбудовування"** математичних моделей в існуючу систему планування та керування;

✓ **труднощі перевірки** у вирішенні нових соціально-економічних задач тощо.

До ефективних засобів подолання цих труднощів можна віднести такі:

✓ **імітаційне моделювання**, що дає змогу керівнику, який приймає рішення, за допомогою ПК включитися у процес побудови економіко-математичної моделі з прийняттям оптимального рішення на її основі (головний принцип імітаційного моделювання: „Що буде, коли...”);

✓ **системний аналіз**, який припускає комплексне проведення дослідження економічних процесів з урахуванням усіх існуючих елементів та взаємозв'язків, вивчення окремих господарських об'єктів як структурних частин більш загальних систем, виявлення ролі кожного з них у функціонуванні економічного процесу в цілому;

✓ **програмно-цільовий метод** планування, оснований на формуванні цілей та підцілей економічного розвитку, на які треба направити найбільші сили і засоби, та розробці програм їх досягнення.

1.4. Класифікація економіко-математичних моделей

Економіко-математичні моделі можна класифікувати за **такими ознаками**:

- 1) призначенням;
- 2) ступенем ймовірності;
- 3) способом опису;
- 4) способом обліку змінювання процесу за часом;
- 5) точністю математичного відображення явищ, що розглядаються.

За **призначенням** моделі доцільно розбити на чотири класи: *імітаційні, балансові, сіткові, оптимізаційні*.

За **ступенем ймовірності** моделі поділяються на два типи: *ймовірні (стохастичні)*, параметри яких та зовнішні зміни мають випадковий характер, *детерміновані*, в яких ігнорується випадковий характер зміни параметрів.

За способом опису моделі діляться на три класи: *аналітичні*, в яких показники описуються математичними формулами або системою формул, *економетричні (статистичні)*, які призначенні для аналізу і прогнозування економічних явищ, що розглядаються, в умовах невизначеності вхідних даних і реалізуються методами математичної статистики, *змішані*, в яких найбільш прості блоки описуються аналітичними залежностями, а в інших блоках, де опис аналітичними формулами може призвести до значних викривлень, використовується економетричне моделювання.

За способом обліку змінювання процесу за часом моделі поділяються на три класи: *статичні*, у яких передбачається, що вхідні параметри не змінюються за часом, *багатокрокові*, у яких час проходження процесу ділиться на "кроки" (інтервали) і в рамках одного кроку процес розглядається статичним, *динамічні*, де враховується безперервна зміна часу.

За точністю математичного відображення явищ, що розглядаються, моделі діляться на дві групи: *лінійні*, залежності в яких мають змінні першого степеня та не включають їх обернених величин і добуток змінних, *нелінійні*.

1.5. Етапи економіко-математичного моделювання

Процес побудови економіко-математичних моделей загального типу складається з таких взаємозв'язаних етапів:

Перший етап – постановка задачі, де формується мета запланованого заходу, ставляться задачі дослідження, проводиться якісний опис об'єкта.

Другий етап – розробка описувальної моделі, де формулюються та обґрунтовуються показники і система основних припущень.

Третій етап – розробка математичної моделі об'єкта, що вивчається, з вибором методів дослідження, програмного забезпечення ПК або складання алгоритму та програми для ПК за новими задачами.

Четвертий етап – розв'язання задачі на базі розробленої моделі, яке складається з реалізації пакета прикладних або розроблених програм для ПК.

П'ятий етап – перевірка та підстроювання моделі, тобто встановлення відповідності моделі описаному економічному процесу.

Шостий етап – надання результатів розв'язання у формі, зручній для вивчення, аналіз матеріалів моделі на основі обробки результатів.

Контрольні запитання

1. Предмет, мета, завдання та основні поняття економіко-математичного моделювання.
2. Історія розвитку економіко-математичних методів.
3. Сучасний стан економіко-математичного моделювання.
4. Класифікація економіко-математичних моделей.
5. Етапи економіко-математичного моделювання.

2. ЗАДАЧІ ЛІНІЙНОГО ПРОГРАМУВАННЯ

Лінійне програмування (ЛП) є одним з розділів математичного програмування.

ЛП – це наука про методи дослідження та відшукування оптимальних значень лінійної функції, на невідомі якої накладені лінійні обмеження.

2.1. Модель загальної задачі лінійного програмування

Загальна задача ЛП полягає в знаходженні екстремуму (максимуму або мінімуму) лінійної цільової функції при наявності обмежень на n змінних у вигляді m лінійних нерівностей або рівнянь та умов невід’ємності змінних. Економіко-математична модель загальної задачі ЛП формулюється таким чином.

Дана система m лінійних рівнянь і нерівностей з n змінними

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2 \\ \dots \\ a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kn}x_n \leq b_k \\ a_{k+1,1}x_1 + a_{k+1,2}x_2 + \dots + a_{k+1,n}x_n = b_{k+1} \\ a_{k+2,1}x_1 + a_{k+2,2}x_2 + \dots + a_{k+2,n}x_n = b_{k+2} \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (2.1)$$

і лінійна функція

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (2.2)$$

Необхідно знайти такий розв’язок системи $X = (x_1, x_2, \dots, x_j, \dots, x_n)$, де

$$x_j \geq 0 (j = 1, 2, \dots, l; \quad l \leq n), \quad (2.3)$$

при якому лінійна функція F (2.2) набуває оптимального (тобто максимального або мінімального) значення.

Система (2.1) називається *системою обмежень*, а функція F – *лінійною функцією, лінійною формою, цільовою функцією або функцією мети*.

Більш коротко загальну задачу лінійного програмування можна подати у вигляді

$$F = \sum_{j=1}^n c_j x_j \rightarrow \max \quad (\text{або } \rightarrow \min)$$

при обмеженнях:

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq b_i & (i = 1, 2, \dots, k) \\ \sum_{j=1}^n a_{ij} x_j = b_i & (i = k + 1, k + 2, \dots, m) \\ x_j \geq 0 (j = 1, 2, \dots, n; \quad l \leq n) \end{cases}$$

Оптимальним розв'язком (або *оптимальним планом*) задачі лінійного програмування називається розв'язок $X = (x_1, x_2, \dots, x_j, \dots, x_n)$ системи обмежень (2.1), що задовольняє умову (2.3), при якому лінійна функція (2.2) набуває оптимального (максимального або мінімального) значення.

Терміни "розв'язок" і "план" – синоніми, однак перший використовується частіше, коли мова йде про формальну сторону задачі (її математичний розв'язок), а другий – про змістовну сторону (економічну інтерпретацію).

За умови, що всі змінні невід'ємні $x_j \geq 0 (j = 1, 2, \dots, n)$, система обмежень (2.1) складається лише з одних нерівностей, така задача лінійного програмування називається *стандартною*, якщо система обмежень складається з одних рівнянь, то задача називається *канонічною*. Будь-яка задача лінійного програмування може бути зведена до канонічної, стандартної або загальної задачі. Розглянемо спочатку допоміжну теорему.

Теорема 2.1. *Усякому розв'язку нерівності $(\alpha_1, \alpha_2, \dots, \alpha_n)$*

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \dots + \alpha_{in}x_n \leq b_i \quad (2.4)$$

відповідає певний розв'язок $(\alpha_1, \alpha_2, \dots, \alpha_n; \alpha_{n+i})$ рівняння

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \dots + \alpha_{in}x_n + x_{n+i} = b_i, \quad (2.5)$$

$$\text{у якому} \quad x_{n+i} \geq 0, \quad (2.6)$$

і, навпаки, кожному розв'язку $(\alpha_1, \alpha_2, \dots, \alpha_n; \alpha_{n+i})$ рівняння (2.5) і нерівності (2.6) відповідає певний розв'язок $(\alpha_1, \alpha_2, \dots, \alpha_n)$ нерівності (2.4).

Зауваження. У розглянутій теоремі всі нерівності вигляду " \leq ", тому додаткові невід'ємні змінні вводяться зі знаком "+". У випадку нерівності вигляду " \geq " додаткові змінні варто було б увести зі знаком "-".

Існує ще 2 види запису канонічної задачі: матрична та векторна.

Матрична форма запису:

$$F = CX \rightarrow \max \quad (\min) \quad (2.7)$$

При обмеженнях

$$AX = B \quad (2.8)$$

$$X \geq 0, \quad (2.9)$$

$$\text{де } C = (c_1, c_2, \dots, c_n); A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}; X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}; B = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix}.$$

Тут C – матриця-рядок, A – матриця системи, X – матриця-стовпець змінних, B – матриця-стовпець вільних членів.

Векторна форма запису:

$$F = CX \rightarrow \max \quad (\min) \quad (2.10)$$

При обмеженнях

$$P_1 x_1 + P_2 x_2 + \dots + P_n x_n = P \quad (2.11)$$

$$X \geq 0, \quad (2.12)$$

де CX – скалярний добуток векторів $C = (c_1, c_2, \dots, c_n)$ і $X = (x_1, x_2, \dots, x_n)$, вектори P_1, P_2, P_n, P складаються відповідно з коефіцієнтів при змінних і вільних членів:

$$P_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \quad P_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \dots, \quad P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}, \quad P = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}.$$

Векторна нерівність $X \geq 0$ означає, що всі компоненти вектора X невід'ємні, тобто $x_j \geq 0, j = 1, 2, \dots, n$.

Скалярним добутком двох векторів $C = (c_1, c_2, \dots, c_n)$ і $X = (x_1, x_2, \dots, x_n)$ називається число, що дорівнює сумі добутків відповідних координат цих векторів, тобто $CX = c_1x_1 + c_2x_2 + \dots + c_nx_n$.

Правила зведення будь-якої лінійної задачі до стандартного вигляду:

1. Від задачі на максимум функції мети до еквівалентної задачі на мінімум можна перейти, використовуючи формулу $\max Z = -\min(-Z)$.
2. Якщо права частина основного обмеження негативна, то це обмеження слід помножити на (-1) .
3. Нерівності в еквівалентні рівняння слід перетворювати за допомогою додаткових змінних. Невід'ємна змінна називається додатковою, якщо вона з коефіцієнтом $+1$ або -1 уводиться в ліву частину нерівності.

Невід'ємна змінна називається додатковою, якщо вона з коефіцієнтом $+1$ або -1 уводиться в ліву частину нерівності для перетворення його в еквівалентне рівняння. Обмеження-нерівність початкової задачі лінійного програмування, яке має вигляд \leq , можна перетворювати в обмеження-рівняння, додаючи до його лівої частини додаткову невід'ємну змінну, а обмеження-нерівність типу \geq – в обмеження-рівняння, віднімаючи з його лівої частини додаткову невід'ємну змінну. Число введених змінних дорівнює числу перетворених нерівностей у рівняння.

2.2. Геометрична інтерпретація задачі лінійного програмування.

Елементи геометрії опуклих множин

Множина точок називається опуклою, якщо вона разом з будь-якими двома своїми точками містить увесь відрізок, що з'єднує ці точки. Згідно із цим визначенням багатокутник на рис. 2.1, а є опуклою множиною, а багатокутник на рис. 2.1, б не є таким, оскільки відрізок MN між двома його точками M і N не повністю належить цьому багатокутнику.

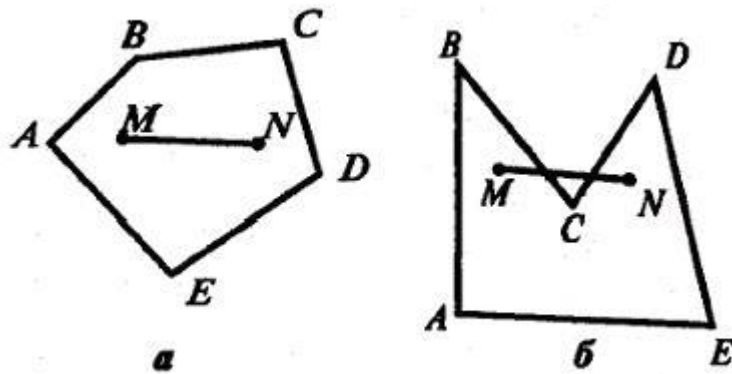


Рисунок 2.1

Прикладами опуклих множин є коло, сектор, відрізок, багатокутна область, куб, піраміда і т.д. Опуклі множини мають властивість, яка встановлюється наступною теоремою.

Теорема 2.2.1 Перетин (спільна частина) будь-якого числа опуклих множин є опукла множина.

Доведення (для доведення теореми обмежимося випадком двох множин):

Нехай M і N – будь-які дві точки перетину двох множин A і B (рис. 2.2).

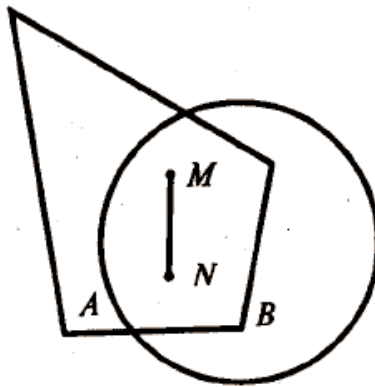


Рисунок 2.2

Оскільки точки M і N належать перетину множин, тобто одночасно опуклій множині A і опуклій множині B , то згідно з визначенням опуклої множини всі точки відрізка MN будуть належати як множині A , так і множині B , тобто перетину цих множин. А це й означає, що перетин даних множин і є опуклою множиною.

Серед точок опуклої множини можна виділити внутрішні, межові й кутові точки.

Точка множини називається *внутрішньою* (точка M на рис. 2.3), якщо в деякому її околі містяться точки тільки даної множини. Під околom точки (простору) мають на увазі коло (куля) із центром у цій точці.

Точка множини називається *межовою* (точка N на рис. 2.3), якщо в будь-якому її околі містяться як точки, що належать даній множині, так і точки, що не належать їй.

Точка множини називається *кутовою* (або *крайньою*) (точки A, B, C, D, E на рис. 2.3), якщо вона не є внутрішньою для жодного відрізка, що цілком належить даній множині. Для опуклої множини кутові точки завжди збігаються з вершинами багатокутника, тоді як для неопуклої множини це необов'язково.

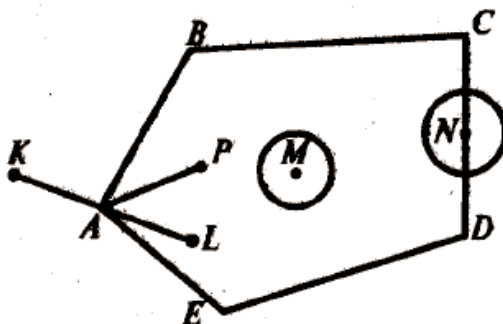


Рисунок 2.3

Множина точок називається *замкненою*, якщо вона включає всі свої межові точки. Множина точок називається *обмеженою*, якщо існує куля (коло) радіуса кінцевої довжини із центром у будь-якій точці множини, який повністю містить у собі дану множину, а якщо ні, то множина називається *необмеженою*.

Якщо фігура обмежена тільки прямими або їх відрізками, то число її кутових точок обмежене; у випадку криволінійності меж фігура містить нескінченно багато кутових точок.

Аналітично точка опуклої множини зображується впорядкованою парою чисел (x_1, x_2) або впорядкованою трійкою чисел (x_1, x_2, x_3) . Поняття точки можна узагальнити, припускаючи під *точкою* (або *вектором*) упорядкований набір n чисел $X = (x_1, x_2, \dots, x_n)$, у якому числа x_1, x_2, \dots, x_n називаються *координатами точки (вектора)*. Наприклад, для характеристики

Множина усіх точок $X = (x_1, x_2, \dots, x_n)$ утворює n -вимірний точковий (векторний) простір.

2.3. Приклади задач лінійного програмування і сформованих на їх основі оптимізаційних моделей

Позначимо x_j ($j = 1, 2, \dots, n$) – число одиниць продукції P_j , запланованої до виробництва, b_i ($i = 1, 2, \dots, n$) – запас ресурсу S_i , a_{ij} – число одиниць ресурсу S_i , що витрачається на виготовлення одиниці продукції P_j (числа a_{ij} часто називають технологічними коефіцієнтами), c_j – прибуток від реалізації одиниці продукції P_j .

[illegible]
$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0, \quad (2.14)$$

при якому функція

$$F = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (2.15)$$

набуває максимального значення.

2.3.2. Задача складання раціону (задача про дієту, задача про суміші). Нехай при відгодівлі тварин використовується n видів кормів, що містять m живильних речовин. Позначимо: x_j ($j = 1, 2, \dots, n$) – число одиниць корму n -го виду, b_i ($i = 1, 2, \dots, m$), – необхідний мінімум змісту в раціоні живильної речовини S_i , a_{ij} – число одиниць живильної речовини S_i в одиниці корму j -го виду, c_j – вартість одиниці корму j -го виду. Тоді економіко-математична модель задачі набуває вигляду:

знайти такий раціон $X = (x_1, x_2, \dots, x_n)$, що задовольняє системі

[illegible]

i умові

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0, \quad (2.17)$$

при якому функція

$$F = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (2.18)$$

набуває максимального значення.

2.3.3. Задача про використання потужностей (задача про завантаження встаткування). Підприємству заданий план виробництва продукції за часом і номенклатурою: потрібно за час T випустити n_1, n_2, \dots, n_k одиниць продукції P_1, P_2, \dots, P_k . Продукція виробляється на верстатах S_1, S_2, \dots, S_m . Для кожного верстата відомі продуктивність a_{ij} (тобто число одиниць продукції P_j , яке можна зробити на верстаті S_i) і витрати b_{ij} на виготовлення продукції P_j на верстаті S_i в одиницю часу.

Необхідно скласти такий план роботи верстатів (тобто так розподілити випуск продукції між верстатами), щоб витрати на виробництво всієї продукції були мінімальними.

Розв'язання. Складемо економіко-математичну модель задачі.

Позначимо x_{ij} – час, протягом якого верстат S_i буде зайнятий виготовленням продукції P_j ($i = 1, 2, \dots, m, j = 1, 2, \dots, k$).

правдиві нерівності:

(2.19)

валися наступні рівняння:

(2.20)

Крім того,

(2.21)

Витрати на виробництво всієї продукції виражаться функцією

$$(2.22)$$

буде вигляду: знайти такий розв'язок $X = (x_{11}, x_{12}, \dots, x_{mk})$, що задовольняє системам (2.19) і (2.20) і умові (2.21), при якому функція (2.22) набуває мінімального значення.

2.3.4. Задача про розкрій матеріалів

сті a одиниць. Потрібно виготовити з нього l різних комплектуючих виробів у кількостях, пропорційних числам $b_1, b_2 \dots b_l$ (умова комплектності). Кожна одиниця матеріалу може бути розкроєна n різними способами, причому використання i -го способу ($i = 1, 2, \dots, n$) дає a_{ik} одиниць k -го виробу ($k = 1, 2, \dots, l$).

*mi*ϑ.

Складемо економіко-математичну модель задачі.

— ЧИСЛО КОМПЛЕКТІВ ВИРОБІВ, ЩО ВИГОТОВЛЯЮТЬСЯ.

Оскільки загальна кількість матеріалу дорівнює сумі його одиниць, що розкрояють різними способами, то

$$\sum_{i=1}^n x_i = a. \quad (2.23)$$

Вимога комплектності виразиться рівняннями

$$\sum_{i=1}^n x_i a_{ik} = b_k x. \quad (k = 1, 2, \dots, l). \quad (2.24)$$

Очевидно, що

$$x_i \geq 0 \quad (i = 1, 2, \dots, n). \quad (2.25)$$

Економіко-математична модель задачі: *знайти такий розв'язок $X = (x_1, x_2, \dots, x_n)$, що задовольняє системі рівнянь (2.23) — (2.24) і умові (2.25), при якому функція $F = x$ набуває максимального значення.*

Завдання про розкрій легко узагальнити на випадок m матеріалів, що розкрояються.

Нехай кожна одиниця j -го матеріалу ($j = 1, 2, \dots, m$) може бути розкrojена n різними способами, причому використання i -го способу ($i = 1, 2, \dots, n$) дає a_{ijk} одиниць k -го виробу ($k = 1, 2, \dots, l$), а запас j -го матеріалу дорівнює a_j одиниць.

Позначимо x_{ij} — число одиниць j -го матеріалу, що розкривається i -м способом.

Економіко-математична модель задачі про розкрій у загальній постановці набуває вигляду: знайти такий розв'язок $X = (x_{11}, x_{12}, \dots, x_{nm})$, що задовольняє системі

$$\begin{cases} \sum_{i=1}^n x_{ij} \leq a_j & (j = 1, 2, \dots, m) \\ \sum_{i=1}^n \sum_{j=1}^m x_{ij} a_{ijk} = b_k x & (k = 1, 2, \dots, l) \end{cases}$$

і умові $x_{ij} \geq 0$, при якій функція $F = x$ набуває максимального значення.

2.3.5. Транспортна задача

Є деякий товар, який знаходиться на m складах. Позначимо склади

$$\boxed{b_1} \quad \boxed{b_2} \quad \dots \quad \boxed{b_m}$$

Кількість товарів на j -му складі b_j . Є n магазинів, потреби магазинів у товарі – d_i :

$$\boxed{d_1} \quad \boxed{d_2} \quad \dots \quad \boxed{d_n}$$

c_{ij} – вартість перевезення одиниці товару з j -го складу в i -й магазин;

x_{ij} – кількість перевезеної продукції з j -го складу i -й магазин.

Необхідно організувати план перевезень так, щоб мінімізувати транспортні витрати, тобто

$$\min f(x) = \min \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij}.$$

При обмеженнях $x_{ij} \geq 0, \quad i = \overline{1, n}; j = \overline{1, m};$

$\sum_{i=1}^n x_{ij} \leq b_j$ – з j -го складу не можна вивезти більше товару, ніж там знаходиться;

$\sum_{j=1}^m x_{ij} \geq d_i$ – всі потреби пунктів споживання повинні бути задовільнені.

Контрольні запитання

1. Загальна, стандартна та канонічна задача лінійного програмування.
2. Матрична та векторна форма запису канонічної задачі.
3. Правила зведення будь-якої лінійної задачі до стандартного вигляду.
4. Визначення основних понять геометрії опуклих множин.

5. Формулювання та математична модель задачі про використання ресурсів.
6. Формулювання та математична модель задачі складання раціону.
7. Формулювання та математична модель задачі про використання потужностей.
8. Формулювання та математична модель задачі про розкрій матеріалів.
9. Формулювання та математична модель транспортної задачі.

3. МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ

3.1. Графічний метод

Графічний метод розв'язування задач лінійного програмування базується на їхній геометричній інтерпретації й аналітичних властивостях і застосовується, як правило, при розв'язуванні ЗЛП при $n = 2$ і в окремих випадках при $n = 3$, оскільки дуже важко побудувати багатогранник розв'язків, що утвориться в результаті перетину напівплощин. ЗЛП при $n > 3$ зобразити геометрично взагалі неможливо.

Розглянемо ЗЛП при $n = 2$ і розв'яжемо її графічним методом.

Знайти екстремум (max, min) функції:

$$F = c_1x_1 + c_2x_2 \longrightarrow \max (\min) \quad (3.1)$$

$$\text{за умов} \quad \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 \{\leq, =, \geq\} b_1, \\ a_{21}x_1 + a_{22}x_2 \{\leq, =, \geq\} b_2, \\ \\ a_{m1}x_1 + a_{m2}x_2 \{\leq, =, \geq\} b_m, \end{array} \right. \quad (3.2)$$

$$x_1 \geq 0, x_2 \geq 0. \quad (3.3)$$

Припустимо, що система (3.2) за умов (3.3) сумісна й багатокутник її розв'язків обмежений. Відповідно до геометричної інтерпретації ЗЛП кожне i -те обмеження-нерівність (3.2) визначає напівплощину з межевою прямою $a_{i1}x_1 + a_{i2}x_2 = b_i$ ($i = \overline{1, m}$). Системою обмежень (3.2) описується спільна частина або перетину всіх зазначених напівплощин, тобто множина точок, координати яких задовольняють всім обмеженням (3.2).

Умова (3.3) невід'ємності змінних означає, що область допустимих розв'язків задачі належить першому квадранту системи координат двовимірного простору. Цільова функція ЗЛП геометрично інтерпретується як сімейство паралельних прямих $c_1x_1 + c_2x_2 = \text{const}$.

З урахуванням властивостей розв'язків ЗЛП розв'язати ЗЛП графічно означає знайти таку вершину багатокутника розв'язків, у результаті підстановки координат якої в (3.1) лінійна цільова функція набуде \max (\min).

Алгоритм графічного методу розв'язування ЗЛП

1. Побудуємо прямі лінії, рівняння яких одержимо заміною в обмеженнях (3.2) знаків нерівностей на знаки рівностей.
2. Визначимо напівплощини, що відповідають кожному обмеженню задачі.
3. Знайдемо багатокутник розв'язків ЗЛП.
4. Побудуємо вектор $\bar{N} = (c_1; c_2)$, що задає напрямок зростання значень цільової функції задачі.
5. Побудуємо пряму $c_1x_1 + c_2x_2 = \text{const}$, перпендикулярну до вектора $\bar{N} = (c_1; c_2)$.
6. Переміщуючи пряму $c_1x_1 + c_2x_2 = \text{const}$ у напрямку вектора $\bar{N} = (c_1; c_2)$ (для задачі максимізації) або в протилежному напрямку (для задачі мінімізації), знайдемо вершину багатокутника розв'язків (останню спільну точку графіка цільової функції й ОДР), де цільова функція досягає екстремального значення.
7. Визначимо координати точки, у якій цільова функція досягає максимального (мінімального) значення, й обчислимо екстремальне значення цільової функції в цій точці.

Переваги графічного методу розв'язування задач лінійного програмування полягають у тому, що він простий, наочний, дозволяє швидко й легко одержати відповідь.

Проте він має наступні **недоліки**:

1. Можливі "технічні" похибки, які неминуче виникають при наближеній побудові графіків.
2. Багато величин, що мають чіткий економічний зміст (такі, як залишки ресурсів виробництва, надлишок живильних речовин і т.п.), не виявляються при графічному розв'язуванні задач.
3. Саме головне: графічний метод неприйнятний для розв'язування практичних задач. Його можна застосувати тільки в тому випадку, коли число змінних у стандартній задачі дорівнює двом. Тому необхідні аналітичні методи, що дозволяють розв'язувати задачі лінійного програмування з будь-яким числом змінних і виявити економічний зміст вхідних у них величин.

3.2. Симплекс-метод

3.2.1. Зведення загальної ЗЛП до канонічної форми

Графічний метод для визначення оптимального плану задачі лінійного програмування доцільно застосовувати, як правило, тільки для задач із двома змінними. При більшій кількості змінних звертаються до загального методу розв'язання задач лінійного програмування – так званого **симплекса-методу**. Слід зауважити, що останній адаптований до розв'язування ЗЛП у канонічній формі. У випадку, якщо ЗЛП подана не в канонічній формі, для застосування симплекса-методу необхідно попередньо її звести до такої форми ЗЛП. Здійснюється це наступним чином.

Якщо i -те обмеження-нерівність у правій частині має значення $b_i < 0$, то, помноживши i -те обмеження на (-1) , у правій частині утвориться додатне значення.

Якщо i -те обмеження має вигляд нерівності

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i,$$

то, ввівши додаткову змінну $x_{n+1} \geq 0$, обмеження-нерівність можна подати у вигляді рівняння

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n + x_{n+1} = b_i.$$

Аналогічно, якщо i -те обмеження має вигляд нерівності

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \geq b_i,$$

то, ввівши додаткову змінну $x_{n+1} \geq 0$, обмеження-нерівність можна подати у вигляді рівняння

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n - x_{n+1} = b_i.$$

Зауваження

Від задачі на максимум цільової функції до еквівалентної задачі на мінімум можна перейти, використовуючи формулу $\max Z = -\min(-Z)$.

3.2.2. Симплекс-метод розв'язування ЗЛП

Процес розв'язування задачі симплексом-методом має ітераційний характер: обчислювальні процедури (ітерації) одного й того самого типу повторюються в певній послідовності доти, доки не буде отриманий оптимальний план задачі або з'ясовано, що його не існує.

З властивостей розв'язків ЗЛП випливає, що за умови існування оптимальний розв'язок знаходиться в крайній точці багатогранника розв'язків.

Кожній такій точці відповідає опорний план. Кожен опорний план визначається системою m лінійно-незалежних векторів, що знаходяться серед векторів A_1, A_2, \dots, A_n . Для знаходження оптимального плану достатньо досліджувати лише опорні плани. Кількість опорних планів може дорівнювати числу комбінацій без повторень C_m^n . При більших значеннях m і n знайти оптимальний план ЗЛП, перебираючи всі опорні плани, досить важко. Тому необхідно мати таку схему, що дозволяє здійснювати впорядкований перехід від одного опорного плану до іншого. Такою схемою і є **симплекс-метод**, що дозволяє за кінцеву кількість кроків, виходячи з відомого опорного плану, одержати оптимальний план. Кожний із кроків полягає в знаходженні такого опорного плану, у якому значення цільової функції краще, ніж на попередньому кроці.

Отже *симплекс-метод* – це поетапна обчислювальна процедура, в основі якої лежить принцип послідовного поліпшення значень цільової функції за рахунок цілеспрямованого переходу від одного опорного плану ЗЛП до іншого.

Алгоритм розв'язування ЗЛП симплексом-методом складається з наступних етапів.

1. Визначення початкового опорного плану задачі лінійного програмування.
2. Побудова симплексної таблиці.
3. Перевірка опорного плану на оптимальність за допомогою оцінок $Z_j - C_j$. Якщо всі оцінки задовольняють умову оптимальності, то визначений опорний план є оптимальним планом задачі. Якщо хоча б одна з оцінок $Z_j - C_j$ не задовольняє умову оптимальності, то переходять до нового опорного плану або встановлюють, що оптимального плану задачі не існує.
4. Перехід до нового опорного плану задачі виконується визначенням розв'язуючого елемента й розрахунком нової симплексної таблиці.
5. Повторення дій, починаючи з п. 3.

Розглянемо докладніше кожний з етапів алгоритму.

1. Визначення першого опорного плану починають із запису ЗЛП у канонічній формі. Після зведення задачі до канонічного вигляду її записують у векторній формі.

За визначенням опорного плану ЗЛП його утворюють m одиничних лінійно-незалежних векторів, які становлять базис m -вимірного простору (де m – кількість обмежень у ЗЛП).

На цьому етапі розв'язування задачі можливі такі випадки:

✓ після запису задачі у векторній формі в системі обмежень є необхідна кількість одиничних векторів. Тоді початковий опорний план визначається безпосередньо без додаткових дій і являє собою вектор B ;

✓ у системі обмежень немає необхідної кількості одиничних незалежних векторів. Тоді для побудови першого опорного плану застосовують **метод штучного базису**, аналізу якого присвячений п. 3.3.

Визначені одиничні лінійно-незалежні вектори, які утворюють базис, і змінні задачі, що відповідають їм, називають **базисними**. Всі інші змінні – **вільні**, їх прирівнюють до нуля й з кожного обмеження задачі визначають значення базисних змінних. У такий спосіб одержують початковий опорний план задачі лінійного програмування.

2. Подальший обчислювальний процес і перевірку опорного плану на оптимальність подають у вигляді симплексної таблиці.

У першому стовпці i таблиці визначається кількість m базисних змінних.

У другому стовпці таблиці – «Базис» – записують вектори A_j , які знаходяться біля базисних змінних у системі обмежень ЗЛП у векторній формі, причому в тій послідовності, у якій вони розміщуються в системі обмежень задачі.

У третьому стовпці симплексної таблиці — « $C_{\text{баз}}$ » — записуються коефіцієнти при базисних змінних у цільовій функції задачі.

Четвертий стовпець таблиці — опорний план ЗЛП.

В інших стовпцях симплексної таблиці, кількість яких відповідає кількості змінних задачі, записують відповідні коефіцієнти з кожного обмеження ЗЛП.

3. Перевіряють опорний план на оптимальність згідно з наведеною далі теоремою.

3.2.3. Теорема (ознака оптимальності опорного плану)

Опорний план $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ ЗЛП є оптимальним, якщо для всіх j ($j = \overline{1, n}$) виконується умова

$$Z_j - C_j \geq 0 \text{ (для задачі на max)}$$

або

$$Z_j - C_j \leq 0 \text{ (для задачі на min).}$$

Значення оцінок $Z_j - C_j$ визначаються за формулою

$$\Delta_j = Z_j - C_j = \sum_{i=1}^m c_i a_{ij} - C_j \quad (j = \overline{1, n})$$

або безпосередньо із симплексної таблиці як скалярний добуток векторів стовпців « $C_{\text{баз}}$ » й « A_j » мінус відповідний коефіцієнт C_j . Розраховані оцінки записують в окремий рядок симплексної таблиці, який називають **оціночним** й позначають $m + 1$.

У процесі перевірки умови оптимальності можливі такі випадки:

а) усі $\Delta_j \quad (j = \overline{1, n})$ задовольняють умову оптимальності, і тоді визначений опорний план є оптимальним,

б) не всі Δ_j задовольняють умову оптимальності, і тоді потрібно виконати перехід до наступного, нового опорного плану задачі.

4. Перехід від одного опорного плану до іншого виконується зміною базису, тобто виключенням з нього деякої змінної й введенням замість неї нової із числа вільних змінних задачі.

Змінна, котра включається в новий базис, відповідає тій оцінці $Z_j - C_j$, що не задовольняє умову оптимальності. Якщо таких оцінок кілька, то серед них вибирають найбільшу за абсолютною величиною й відповідну їй змінну вводять у базис. Припустимо, що індекс зазначеної змінної $j=k$. Відповідний стовпець симплексної таблиці називають **направляючим**.

Для визначення змінної, яка повинна бути виключена з базису, знаходять для всіх позитивних a_{ik} направляючого стовпця величину $\theta = b_i / a_{ik}$ й записують її в останній стовпець симплексної таблиці. Далі вибирають найменше значення, яке вказує на змінну, що виводиться з базису. Припустимо, що це виконується для $i = r$. Відповідний рядок симплексної таблиці буде називатися **направляючим**. Іноді направляючі стовпець і рядок ще називають **розв'язуючими**. Перетином розв'язуючого стовпця й розв'язуючого рядка визначається число симплексної таблиці a_{rk} , яке називають **розв'язуючим елементом**. За допомогою елемента a_{rk} і методу Жордано–Гаусса розраховують нову симплексну таблицю.

Далі ітераційний процес повторюють доти, доки не буде визначений оптимальний план задачі.

У випадку застосування симплекса-методу для розв'язування ЗЛП можливі наступні випадки:

а) Якщо в оціночному рядку останньої симплексної таблиці оцінка $Z_j - C_j = 0$ відповідає вільній (небазисній) змінній, то це означає, що ЗЛП має альтернативний оптимальний план. Одержати його можна, вибравши розв'язуючий елемент у відзначеному стовпці таблиці й виконавши один крок симплекса-методу.

б) Якщо при переході в методі симплекса від одного опорного плану задачі до іншого в направляючому стовпці немає позитивних елементів a_{rk} , тобто неможливо вибрати змінну, яка повинна бути виведена з базису, то це означає, що цільова функція ЗЛП є необмеженою в даній області й оптимальних планів не існує.

3.3. Метод штучного базису

Розв'язування ЗЛП симплексом-методом починається зі знаходження першого опорного плану. Вище вказувалося, що такий план може бути знайдений шляхом зведення системи рівнянь-обмежень до одиничного базису. Однак існує й інший метод побудови початкового опорного плану. Цей метод може застосовуватися, наприклад, у випадку, коли в системі обмежень не вистачає необхідної кількості одиничних лінійно-незалежних векторів. Для його аналізу разом зі ЗЛП, що будемо називати **вихідною**, розглянемо розширену задачу, яка складена на основі вихідної задачі в такий спосіб. Припускаючи, що $b_i \geq 0 \ (i = \overline{1, m})$, уведемо в кожне рівняння-обмеження по одній позитивній змінній $x_{n+i} \geq 0 \ (i = \overline{1, m})$, які будемо називати **штучними**, а із цільової функції віднімемо суму штучних змінних, помножену на як завгодно велике додане число M (при розв'язанні задачі на \max). У результаті одержимо так звану M -задачу:

знайти

$$F = \sum_{j=1}^n c_j x_j - M \sum_{i=1}^m x_{n+i} \rightarrow \max \quad (3.4)$$

при обмеженнях

$$\sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i \ (i = \overline{1, m}) \quad (3.5)$$

$$x_j \geq 0 \ (\overline{1, n+m}). \quad (3.6)$$

У системі (3.5) змінні x_{n+i} ($i = \overline{1, m}$) утворюють **базис**, що називається **штучним**. При $x_1 = \dots = x_n = 0$ з (3.5) одержимо початковий опорний план $X_0 = (0, \dots, 0, b_1, \dots, b_m)$ M -задачі.

Далі задача розв'язується на основі застосування алгоритму симплекса-методу.

Зауваження. При розв'язанні ЗЛП методом штучного базису штучні змінні необхідно вводити тільки в ті рівняння-обмеження, які не є розв'язаними відносно «природних» базисних змінних.

Як видно з рівняння (3.4), функція F складається із двох доданків $\sum_{j=1}^n c_j x_j = f$ й $M \sum_{i=1}^m x_{n+i}$, тому в симплексних таблицях замість одного оціночного

рядка $m+1$ розглядаються два рядки: $m+1$ -й для f , $m+2$ -й для $M \sum_{i=1}^m x_{n+i}$.

При цьому оцінки $Z_j - C_j$ записуються в обидві рядки відповідним чином й ознака оптимальності перевіряється за обома рядками. За значеннями оцінок $Z_j - C_j$ визначається змінна, яка повинна бути включена в базис. Перетворення в таблиці продовжують доти, доки в базисі не будуть виключені всі штучні змінні. Штучні змінні назад у базис не вводяться. Після виключення з базису всіх штучних змінних процес пошуку оптимального плану триває з використанням тільки першого рядка цільової функції.

Отже, необхідною умовою оптимальності опорного плану є також вимога, щоб у процесі розв'язування задачі всі штучні змінні були виведені з базису й дорівнювали нулю.

Якщо у вихідній задачі необхідно мінімізувати цільову функцію, то в M -задачі цільова функція буде мати вигляд

$$F = \sum_{j=1}^n c_j x_j + M \sum_{i=1}^m x_{n+i} \rightarrow \min, \quad (3.7)$$

тобто у цільовій функції задачі на \min штучні змінні мають коефіцієнти $(+M)$.

Слід зазначити, що якщо для опорного плану ЗЛП всі оцінки $Z_j - C_j$ ($j = \overline{1, n}$) задовольняють умову оптимальності, але при цьому хоча б одна штучна змінна є базисною й має додатне значення, то це означає, що система обмежень задачі несумісна й оптимальних планів такої задачі не існує.

Якщо M -задача не має розв'язку, то й вихідна задача є нерозв'язаною.

Контрольні запитання

1. Геометричний зміст нерівностей, умови невід'ємності змінних та цільової функції.
2. Алгоритм графічного методу розв'язування ЗЛП, його переваги та недоліки.
3. Зведення загальної ЗЛП до канонічної форми.
4. Поняття симплекса-методу та його алгоритм.
5. Ознака оптимальності опорного плану.
6. Метод штучного базису.

$$\text{Вихідна } A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \text{ двоїста } A' = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

5. Число нерівностей у системі обмежень однієї задачі збігається із числом змінних в іншій задачі.

6. Умови невід'ємності змінних є в обох задачах.

Дві задачі лінійного програмування, що мають зазначені властивості, називаються **симетричними взаємно двоїстими задачами**. Надалі для простоти будемо називати просто **двоїстими задачами**.

Виходячи з визначення, можна запропонувати наступні **правила побудови двоїстих задач**.

1. Звести всі нерівності системи обмежень вихідної задачі до одного змісту: якщо у вихідній задачі шукають максимум лінійної функції, то всі нерівності системи обмежень звести до вигляду " \leq ", а якщо мінімум – до вигляду " \geq ". Для нерівностей, у яких дана вимога не виконується, помножити на (-1) .

2. Скласти розширену матрицю системи A_1 , у яку включити матрицю коефіцієнтів при змінних A , стовпець вільних членів системи обмежень і рядок коефіцієнтів при змінних у лінійній функції.

3. Знайти матрицю A'_1 , транспоновану до матриці A_1 .

4. Сформулювати двоїсту задачу на підставі отриманої матриці A'_1 й умови невід'ємності змінних.

Приклад 1. Скласти задачу, двоїсту наступній задачі:

$$Z = -x_1 + 2x_2 \rightarrow \max$$

при обмеженнях:

$$\begin{cases} 2x_1 - x_2 \geq 1, \\ -x_1 + 4x_2 \leq 24, \\ x_1 - x_2 \leq 3, \\ x_1 + x_2 \geq 5, \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

Розв'язання. 1. Оскільки вихідна задача на максимізацію, то зведемо всі нерівності системи обмежень до вигляду " \leq ", для чого обидві частини першої й четвертої нерівностей помножимо на (-1) . Одержимо

$$\begin{cases} -2x_1 + x_2 \leq -1, \\ -x_1 + 4x_2 \leq 24, \\ x_1 - x_2 \leq 3, \\ -x_1 - x_2 \leq -5. \end{cases}$$

2. Складемо розширену матрицю системи

$$A_1 = \begin{pmatrix} -2 & 1 & -1 \\ -1 & 4 & 24 \\ 1 & -1 & 3 \\ -1 & -1 & -5 \\ -1 & 2 & Z \end{pmatrix}.$$

3. Знайдемо матрицю A'_1 , транспоновану до A :

$$A'_1 = \begin{pmatrix} -2 & -1 & 1 & -1 & -1 \\ 1 & 4 & -1 & -1 & 2 \\ -1 & 24 & 3 & -5 & F \end{pmatrix}.$$

4. Сформулюємо двоїсту задачу:

$$F = -y_1 + 24y_2 + 3y_3 - 5y_4 \rightarrow \min$$

при обмеженнях

$$\begin{cases} 2y_1 - y_2 + y_3 - y_4 \geq -1, \\ y_1 + 4y_2 - y_3 - y_4 \geq 2, \end{cases}$$

$$y_1 \geq 0, y_2 \geq 0, \dots, y_4 \geq 0.$$

4.2. Співвідношення двоїстості

Співвідношення двоїстості – це співвідношення між значеннями змінних пари двоїстих задач.

Між прямою й двоїстою задачами лінійного програмування існує тісний взаємозв'язок, що впливає з наведених далі теорем.

Перша теорема двоїстості

Якщо одна із взаємно двоїстих задач має оптимальний розв'язок, то його має й інша, причому оптимальні значення їхніх цільових функцій дорівнюють одне одному

$$Z_{\max} = F_{\min} \text{ або } Z(X^*) = F(Y^*).$$

Якщо ж цільова функція однієї із задач необмежена, то друга задача взагалі не має розв'язків.

Якщо пряма задача лінійного програмування має оптимальний план X^* , визначений методом симплекса, то оптимальний план двоїстої задачі Y^* визначається так:

$$Y^* = \vec{C}_{\text{баз}} D^{-1}. \quad (4.1)$$

У виразі (4.1) $C_{\text{баз}}$ – вектор-рядок, що складається з коефіцієнтів цільової функції прямої задачі при змінних, які є базисними в оптимальному плані. При цьому порядок наступності координат цього вектора визначається порядком запису базисних змінних в останній симплексній таблиці, D^{-1} – матриця, що обернена до матриці D , яка складається з базисних векторів оптимального плану, компоненти яких взяті з початкового опорного плану задачі.

Обернена матриця D^{-1} завжди міститься в останній таблиці симплекса в тих стовпцях, де в першій таблиці знаходилася одинична матриця. За допомогою зазначеного співвідношення під час визначення оптимального плану однієї із взаємно двоїстих задач лінійного програмування знаходять розв'язок іншої задачі.

Друга теорема двоїстості

Якщо в результаті підстановки оптимального плану прямої задачі в систему обмежень цієї задачі i -те обмеження виконується як строга нерівність то відповідний i -й компонент оптимального плану двоїстої задачі дорівнює нулю.

Якщо i -й компонент оптимального плану двоїстої задачі додатний, то відповідне i -те обмеження прямої задачі виконується для оптимального плану як рівняння.

Третя теорема двоїстості

Двоїста оцінка характеризує приріст цільової функції, який обумовлений малими змінами вільного члена відповідного обмеження. Економічний зміст третьої теореми двоїстості полягає в тому, що відповідна позитивна оцін-

ка показує зростання значення цільової функції прямої задачі, якщо запас відповідного дефіцитного ресурсу збільшується на одну одиницю.

4.3. Економічна інтерпретація двоїстої задачі

Теореми двоїстості дозволяють дати економічну інтерпретацію двоїстості. Розглянемо це питання детальніше. Для цього розглянемо пряму задачу як задачу розподілу ресурсів, тобто як задачу планування виробництва.

Для цього випадку кожен коефіцієнт цільової функції прямої задачі $C_j (j = \overline{1, n})$ являє собою величину прибутку, що з'являється від реалізації одиниці j -го виду продукції, $b_i (i = \overline{1, m})$ запаси сировини i -го виду, $a_{ij} (i = \overline{1, m}, j = \overline{1, n})$ – кількість одиниць сировини i -го виду, що використовуються для виготовлення одиниці j -го виду продукції, $x_j (j = \overline{1, n})$ – кількість одиниць j -го виду продукції, яку необхідно виготовити.

Тоді для економічної інтерпретації змінних $y_i (i = \overline{1, m})$ двоїстої задачі досить проаналізувати розмірності величин Z і F .

Оскільки розмірністю величини є гроші (ця величина характеризує прибуток), то з того, що $Z_{\max} = F_{\min}$ випливає, що розмірністю величини $y_i (i = \overline{1, m})$ є $\frac{\text{гроші}}{\text{обсяг ресурсів}}$, тобто цінність i -го ресурсу. Тому іноді двоїсті змінні називають ще тіньовими цінами.

А, отже, з економічної точки зору цільова функція F визначає загальну вартість усіх ресурсів.

За допомогою двоїстих оцінок можна визначити статус кожного ресурсу прямої задачі й рентабельність виготовленої продукції.

Ресурси, що використовуються для виробництва продукції, можна умовно розділити на дефіцитні й недефіцитні залежно від того, повне або часткове їхнє використання передбачене оптимальним планом прямої задачі. Отже, якщо $y_i > 0$, то i -тий ресурс використовується повністю і є дефіцитним. Якщо ж $y_i = 0$, то i -тий ресурс використовується не повністю і є недефіцитним.

Аналіз рентабельності виготовленої продукції виконується за допомогою двоїстих оцінок й обмежень двоїстої задачі.

Ліва частина j -го обмеження являє собою вартість усіх ресурсів, які використовуються для виробництва одиниці i -тої продукції. Якщо вона більше

C_j , то виробництво нерентабельне й тому в прямій ЗЛП $x_j = 0$. Якщо ж вона дорівнює C_j , то виробництво рентабельне й в оптимальному плані $x_j > 0$.

Контрольні запитання

1. Поняття й вид двоїстої ЗЛП.
2. Правила побудови двоїстих задач.
3. Перша теорема двоїстості.
4. Друга та третя теореми двоїстості.
5. Економічна інтерпретація двоїстої задачі.

5. МОДЕЛІ ЦІЛОЧИСЛОВОГО ЛІНІЙНОГО ПРОГРАМУВАННЯ

5.1. Постановка задачі цілочислового програмування

За змістом значної частини економічних задач, що належать до задач лінійного програмування, компоненти розв'язку повинні виражатися в цілих числах, тобто бути цілочисловими. До них належать, наприклад, задачі оптимального використання устаткування, календарного планування, про призначення, оптимального розкрою матеріалів, задача комівояжера та інші.

Задача лінійного цілочислового програмування записується так: *знайти такий розв'язок (план) $X = (x_1, x_2, \dots, x_n)$, при якому лінійна функція*

$$Z = \sum_{j=1}^n c_j x_j \quad (5.1)$$

набуває максимального або мінімального значення при обмеженнях

$$\sum_{i=1}^n a_{ij} x_j = b_i, i = \overline{1, m} \quad (5.2)$$

$$x_j \geq 0, j = 1, 2, \dots, n \dots \quad (5.3)$$

$$x_j - \text{цілі числа.} \quad (5.4)$$

Слід зазначити, що класична транспортна задача й деякі інші задачі транспортного типу "автоматично" забезпечують розв'язання задачі в цілих числах (якщо, звичайно, цілочислові параметри умов). Однак у загальному випадку умова цілочисельності (5.4), що додається до звичайних задач лінійного програмування, істотно ускладнює її розв'язок.

Для розв'язування задач лінійного цілочислового програмування використовуються ряд методів. Найпростіший з них – звичайний метод лінійного програмування. У випадку, якщо компоненти оптимального рішення виявляються нецілочисловими, їх округляють до найближчих цілих чисел. Цей метод застосовують тоді, коли окрема одиниця сукупності становить малу частину об'єму всієї сукупності. У протилежному випадку округлення може призвести до далекого від оптимального цілочисельного розв'язку, тому використовують спеціально розроблені методи.

Методи цілочислової оптимізації можна розділити на **три основні групи**: а) методи відтинання, б) комбінаторні методи, в) наближені методи. Зупинимося докладніше на **методах відтинання**.

5.2. Методи відтинання. Метод Гоморі

Сутність методів відтинання полягає в тому, що спочатку задача розв'язується без умови цілочисельності. Якщо отриманий план цілочисловий, то задача є розв'язаною. У протилежному випадку до обмежень задачі додається нове обмеження, що має такі властивості:

- воно повинне бути лінійним;
- повинне відтинати знайдений оптимальний нецілочисловий план;
- не повинне відтинати жодного цілочислового плану.

Додаткове обмеження, що має зазначені властивості, називається *правильним відтинанням*.

Далі задача розв'язується з урахуванням нового обмеження. Після цього, якщо буде потреба, додається ще одне обмеження й т.д.

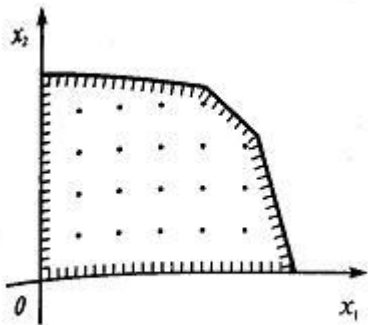


Рисунок 5.1

Геометрично додавання кожного лінійного обмеження відповідає проведенню прямої (гіперплощини), що відтинає від багатокутника (багатогранника) розв'язків деяку його частину разом з оптимальною точкою з нецілими координатами, але не зачіпає ні однієї із цілих точок цього багатогранника. У результаті новий багатогранник розв'язків містить усі цілі точки, які містяться в первісному багатограннику розв'язків і відповідно отриманий при цьому багатограннику оптимальний розв'язок буде цілочисловим (рис. 5.1).

Один з алгоритмів розв'язання задачі лінійного цілочислового програмування (5.1) – (5.4), запропонований Гоморі, оснований на симплексному методі й використовує досить простий спосіб побудови правильного відтинання.

Нехай задача лінійного програмування (5.1)–(5.3) має кінцевий оптимум і на останньому кроці її розв'язання симплексним методом отримані наступні рівняння, що виражають основні змінні $x_1, x_2, \dots, x_j, \dots, x_m$ через неосновні змінні $x_{m+1}, x_{m+2}, \dots, x_{m+i}, \dots, x_n$ оптимального розв'язку

що оптимальним розв'язком задачі (5.1)–(5.3) є $X^* = (\beta_1, \beta_2, \dots, \beta_i, \dots, \beta_m, 0, 0, \dots, 0)$, у якому, наприклад, β_i – нецілий компонент. У цьому випадку можна довести, що нерівність

$$\{\beta_i\} - \{\alpha_{i,m+1}\}x_{m+1} - \dots - \{\alpha_{in}\}x_n \leq 0, \quad (5.6)$$

(У нерівності (5.6) є присутнім символ $\{ \}$, що означає дробову частину числа. Цілою частиною числа a називається найбільше ціле число $[a]$, що не переважає a , а дробовою частиною числа – число $\{a\}$, як різниця між цим числом і його цілою частиною, тобто $\{a\} = a - [a]$. Наприклад, для $a = 2\frac{1}{3} \rightarrow [a] = 2, \{a\} = 2\frac{1}{3} - 2 = \frac{1}{3}$.)

1. Симплексним методом розв'язати задачу (5.1) – (5.3) без урахування умови цілочисельності. Якщо всі компоненти оптимального плану цілі, то він є оптимальним і для задачі цілочислового програмування (5.1) – (5.4). Якщо перша задача (5.1) – (5.3) є нерозв'язаною (тобто не має кінцевого оптимуму або умови її суперечливі), то й друга задача (5.1) – (5.4) також нерозв'язна.

3. Нерівність (5.6) уведенням додаткової невід'ємної цілочислової змінної перетворити в рівносильне рівняння

і включити його в систему обмежень (5.2).

4. Отриману розширену задачу розв'язати симплексним методом. Якщо знайдений оптимальний план буде цілочисловим, то задача цілочислового програмування (5.1) – (5.4) розв'язана. У протилежному випадку повернутися до п. 2 алгоритму.

Якщо задача розв'язна в цілих числах, то після кінцевого числа кроків (ітерацій) оптимальний цілочисловий план буде знайдений.

Якщо в процесі розв'язання з'явиться рівняння (що виражає основну змінну через неосновні) з нецілим вільним членом і цілими іншими коефіцієнтами, то відповідне рівняння не має розв'язку в цілих числах. У цьому випадку й дана задача не має цілочислового оптимального розв'язку.

5.3. Метод гілок і меж

Метод гілок і меж – один з *комбінаторних* методів. Його суть полягає в упорядкованому переборі варіантів і розгляді лише тих з них, які виявляються за певними ознаками перспективними, і відкиданні безперспективних варіантів.

Метод гілок і меж полягає в наступному: множина допустимих розв'язків (планів) деяким способом розбивається на підмножини, кожна з яких цим же способом знову розбивається на підмножини. Процес триває доти, поки не отриманий оптимальний цілочисловий розв'язок вихідної задачі.

Нехай задача 1 (задача (5.1) – (5.3) максимізації лінійної функції Z (без урахування цілочисельності змінних) розв'язана симплексним методом і відомі нижня й верхня межі для кожної цілочислової змінної x_j : $v_j \leq x_j \leq w_j$ ($j=1,2,\dots,n$), а також нижня границя лінійної функції Z_0 , тобто при будь-якому плані X $Z(X) \geq Z_0$. Припустимо для визначеності, що тільки перший компонент x_1^* оптимального плану X^* задачі 1 не задовольняє умові цілочисельності. Тоді з області допустимих розв'язків задачі 1 виключається область: $[x_1^*] < x_1^* < [x_1^*] + 1$, де $[x_1^*]$ – ціла частина числа x_1^* . У результаті із задачі 1 формують дві задачі: 2 і 3, які відрізняються одна від одної тим, що в задачі 2, крім обмежень (5.2) задачі 1, додане обмеження $v_1 \leq x_1^* \leq [x_1^*] + 1$, а в задачі 3, крім тих же обмежень

(5.2), додане обмеження $[x_1^*] + 1 \leq x_1^* \leq w_1$. Одержимо список із двох задач: 2 і 3.

Розв'язуємо одну з них (у будь-якому порядку). Залежно від отриманого розв'язку список задач розширюється або зменшується.

Якщо в результаті розв'язку однієї із задач 2 або 3 отриманий нецілочисловий оптимальний план, для якого $Z(X^*) \leq Z_0$, то дана задача виключається зі списку. Якщо $Z(X^*) > Z_0$, то з даної задачі формуються нові дві задачі.

Якщо отриманий розв'язок X^* задовольняє умові цілочисельності і $Z(X^*) > Z_0$, то значення Z_0 виправляється і за величину Z_0 приймається оптимум лінійної функції отриманого оптимального цілочислового плану. Процес триває доти, поки список задач не буде вичерпаний, тобто всі задачі не будуть розв'язані.

Зауваження 1. Кожна наступна задача, що складається у процесі застосування методу гілок і меж, відрізняється від попередньої лише однією нерівністю – обмеженням. Тому при розв'язуванні кожної наступної задачі не треба розв'язувати її симплексним методом із самого початку (з I кроку). А доцільніше почати розв'язок з *останнього кроку* (ітерації) попередньої задачі, із системи обмежень якої виключити "старі" (одне або два) рівняння-обмеження і ввести в цю систему "нові" рівняння-обмеження.

Зауваження 2. Назва методу гілок і меж пояснюється тим, що в процесі розв'язування задача послідовно "гілкується", замінюючись більш простими. Так, процес розв'язування задачі можна подати у вигляді дерева (схеми), цифри у вузлах (вершинах) якого позначають номери задач (рис. 5.2).

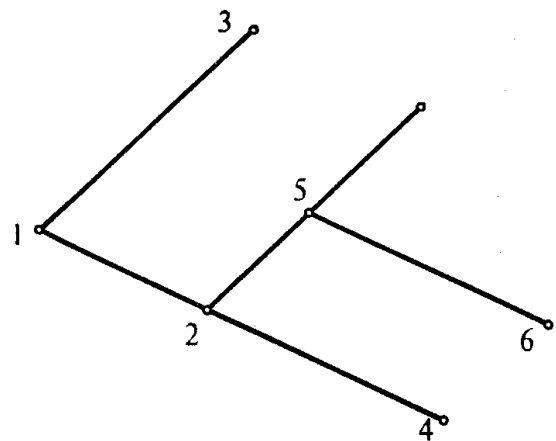


Рисунок 5.2

Контрольні запитання

1. Постановка задачі цілочислового програмування.
2. Сутність методів відтинання.
3. Алгоритм методу Гоморі.
4. Алгоритм методу „гілок і меж”.

6. МОДЕЛІ НЕЛІНІЙНОГО ПРОГРАМУВАННЯ

6.1. Постановка задачі нелінійного програмування

Загальна задача математичного програмування формулюється наступним чином:

знайти вектор $X = (x_1, x_2, \dots, x_n)$, який задовольняє системі обмежень

$$g_i(x_1, x_2, \dots, x_n) \{ \leq, \geq, = \} b_i \quad (i = \overline{1, m}) \quad (6.1)$$

і при якому функція

$$Z = f(x_1, x_2, \dots, x_n) \quad (6.2)$$

набуває екстремального значення.

При цьому передбачається, що функції $g_i(x_1, x_2, \dots, x_n)$ та $f(x_1, x_2, \dots, x_n)$ є відомими.

Звичайно, на деякі змінні x_1, x_2, \dots, x_n накладається умова невід'ємності. Крім того, обмеженням може бути й умова цілочисельності розв'язку для ряду змінних.

Якщо

$$g_i(x_1, x_2, \dots, x_n) = \sum_{j=1}^n a_{ij} x_j \quad (i = \overline{1, m})$$

$$i \quad Z = f(x_1, x_2, \dots, x_n) = \sum_{j=1}^n c_j x_j,$$

де a_{ij}, c_j – відомі константи, то за умови невід'ємності розв'язку задача (6.1) – (6.2) перетворюється в задачу лінійного програмування.

Будь-яку іншу задачу математичного програмування, яка не задовольняє останні умови, будимо називати нелінійною.

Клас задач нелінійного програмування (ЗНП) значно ширше класу задач лінійного програмування. Основні результати в нелінійному програмуванні отримані для випадків, коли система обмежень лінійна, а цільова функція нелінійна. Навіть у таких задачах розв'язок може бути знайдено тільки для вузького класу цільових функцій, а саме для сепарабельних і квадратичних функцій (функція сепарабельна, якщо вона є сумою n функцій $f_j(x_j)$).

Якщо в задачах лінійного програмування точки екстремуму є вершинами багатогранників розв'язків, то в задачах з нелінійною цільовою функцією вони можуть лежати всередині області, на ребрі (грані) або у вершині

багатогранника. Ще більші труднощі виникають при розв'язуванні задач із нелінійними обмеженнями.

Методи нелінійного програмування бувають прямі й непрямі. Прямими методами оптимальні розв'язки шукають у напрямку найшвидшого збільшення (зменшення) цільової функції. Типовими для цієї групи методів є градієнтні. Непрямі методи полягають у зведенні задачі до такої, знаходження оптимуму якої вдається спростити. До них належать найбільш розроблені методи квадратичного й сепарабельного програмування.

Оптимізаційні задачі, на змінні яких накладаються обмеження, розв'язують методами класичної математики. Оптимізацію з обмеженнями - рівняннями виконують методами зведеного градієнта, а саме методом Якобі й множників Лагранжа. У задачах оптимізації з обмеженнями-нерівностями досліджують необхідні й достатні умови існування екстремуму Куна-Таккера.

6.2 Графічний метод розв'язування задач нелінійного програмування

Розгляд задач нелінійного програмування починають із класичної задачі оптимізації. Задачі такого роду мають місце, якщо в обмеженнях є тільки рівняння, відсутні умови невід'ємності та цілочисельності змінних, а функції $g_i(x_1, x_2, \dots, x_n)$ та $f(x_1, x_2, \dots, x_n)$ безперервні й мають часткові похідні не нижче другого порядку. Класичні методи оптимізації при цьому є теоретичним апаратом, що дозволяє в ряді випадків обґрунтувати розробку відповідного обчислювального методу.

Одним з найпростіших типів ЗНП є ЗНП із двома змінними. Такі задачі мають такий вигляд.

Знайти

$$Z = f(x_1, x_2) \rightarrow \max(\min) \quad (6.3)$$

при

$$g_i(x_1, x_2) \{ \leq, \geq, = \} b_i \quad (i = \overline{1, m}) \quad (6.4)$$

$$x_1, x_2 \geq 0. \quad (6.5)$$

Для їх розв'язання можна застосувати **графічний метод**. Суть його в наступному.

1. Будуємо ОДР, що визначається обмеженнями (6.4) – (6.5).
2. Будуємо лінії рівня, які визначаються цільовою функцією (6.3).
3. Установлюються точки, що підозрілі на екстремум. Такими точками є точки, що належать і графіку цільової функції, і ОДР, і які в достатньо малому своєму околі інших точок ОДР не містять.
4. Знаходяться значення цільової функції (6.3) у знайдених точках і серед них вибираються ті, в яких досягається найбільше й найменше значення. В окремих випадках на основі застосування достатніх умов екстремуму визначаються точки екстремуму.

6.1. Метод множників Лагранжа

У математичному аналізі задачі типу: знайти значення змінних x_1, x_2, \dots, x_n , які задовольняють системі обмежень

$$\begin{aligned}\varphi_i(x_1, x_2, \dots, x_n) &= b_i \quad (i = \overline{1, k}), \\ \varphi_i(x_1, x_2, \dots, x_n) &\leq b_i \quad (i = \overline{k+1, m})\end{aligned}\tag{6.6}$$

$$\text{і забезпечують екстремум функції } Z = f(x_1, x_2, \dots, x_n),\tag{6.7}$$

називають задачами на умовний екстремум.

Якщо серед обмежень немає нерівностей й умов невід'ємності або дискретності змінних, $m < n$, функції f й g_i неперервні разом із частинними похідними принаймні до другого порядку включно, то наведена задача на умовний екстремум зводиться до задачі:

$$Z = f(x_1, x_2, \dots, x_n) \rightarrow \max(\min)\tag{6.8}$$

$$\varphi_i(x_1, x_2, \dots, x_n) = b_i \quad (i = \overline{1, m})\tag{6.9}$$

Ці задачі нелінійного програмування називають *класичними задачами оптимізації*. Їх у деяких випадках можна розв'язувати методами диференціального числення.

У найпростішому випадку умовним екстремумом функції $f(x_1, x_2)$ називається максимум або мінімум цієї функції, досягнутий за умови, що x_1 й x_2 задовольняють додаткову умову (рівнянню зв'язку)

$$\varphi(x_1, x_2) = b.$$

Умовний екстремум функції $f(x_1, x_2)$ при існуванні додаткового обмеження $\varphi(x_1, x_2) = b$ знаходять за допомогою функції Лагранжа

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda[b - \varphi(x_1, x_2)],$$

де λ – невід’ємний сталий множник (множник Лагранжа), безумовний екстремум якого збігається з умовним екстремумом даної функції $f(x_1, x_2)$. Пояснюється це тим, що для точок (x_1, x_2) , які задовольняють умову $\varphi(x_1, x_2) = b$, другий доданок згортається в нуль, і тоді $L = f$. Для інших точок $L \neq f$. Звідси і випливає, що задача про знаходження умовного екстремуму функції $f(x_1, x_2)$ може бути замінена знаходженням звичайного екстремуму функції L , оскільки в області допустимих розв’язків функцію $f(x_1, x_2)$ можна замінити функцією Лагранжа.

Необхідна умова екстремуму зводиться до існування розв’язку системи трьох рівнянь із трьома невідомими

$$\begin{cases} \frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} - \lambda \frac{\partial \varphi}{\partial x_1} = 0, \\ \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} - \lambda \frac{\partial \varphi}{\partial x_2} = 0, \\ \frac{\partial L}{\partial \lambda} = b - \varphi(x_1, x_2) = 0. \end{cases}$$

Є й достатні умови, при виконанні яких розв’язок (x_1, x_2, λ) системи визначає точку (стаціонарну), у якій $f(x_1, x_2)$ досягає екстремуму. Це питання розв’язується на підставі вивчення знака диференціала другого порядку d^2L . Оскільки в стаціонарній точці повний диференціал функції $\varphi(x_1, x_2)$ дорівнює нулю, тобто

$$\frac{\partial \varphi}{\partial x_1} dx_1 + \frac{\partial \varphi}{\partial x_2} dx_2 = 0, (dx_1^2 + dx_2^2 \neq 0),$$

і, крім того, $\frac{\partial^2 L}{\partial \lambda^2} = 0$, то диференціал 2-го порядку функції L дорівнює

$$d^2 L = \frac{\partial^2 L}{\partial x_1^2} dx_1^2 + 2 \frac{\partial^2 L}{\partial x_1 \partial x_2} dx_1 dx_2 + \frac{\partial^2 L}{\partial x_2^2} dx_2^2.$$

Функція $f(x_1, x_2)$ в стаціонарній точці (x_1, x_2, λ) має умовний максимум, якщо в ній $d^2 L < 0$, і умовний мінімум, якщо $d^2 L > 0$.

Аналогічно знаходиться умовний екстремум функції трьох і більше змінних при існуванні одного або декількох додаткових обмежень (кількість яких повинна бути менше кількості змінних).

Ідея методу множників Лагранжа розв'язування задачі (6.8) – (6.9) у загальному випадку полягає в заміні даної задачі на просту: на знаходження екстремуму складної функції, але без обмежень.

Ця функція називається функцією Лагранжа і задається у вигляді

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i [b_i - \varphi_i(x_1, \dots, x_n)], \quad (6.10)$$

де λ_i ($i = \overline{1, m}$) – множники Лагранжа.

Алгоритм методу множників Лагранжа може бути поданий у наступному вигляді.

1. Для задачі (6.8) – (6.9) формується функція Лагранжа (6.10).
2. Формується система рівнянь

$$\begin{cases} \frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial \varphi_i}{\partial x_j} = 0 \quad (j = \overline{1, n}), \\ \frac{\partial L}{\partial \lambda_i} = b_i - \varphi_i(x_1, \dots, x_n) = 0 \quad (i = \overline{1, m}). \end{cases} \quad (6.11)$$

3. У результаті розв'язання системи (6.11) знаходяться всі стаціонарні точки $X^* = (x_1, x_2, \dots, x_n)$ й $\lambda^* = (\lambda_1, \lambda_2, \dots, \lambda_n)$. Оскільки вони визначені з необхідної умови екстремуму, то в них можливо досягається максимум або мінімум.

4. Серед стаціонарних точок $X^* = (x_1, x_2, \dots, x_n)$ функції L вибираються такі, у яких функція f має умовні екстремуми при існуванні обмежень. Цей вибір здійснюється, наприклад, за допомогою достатніх умов екстремуму.

6.2. Опуклі та вгнуті функції

Функція $F(X) = F(x_1, \dots, x_n)$, що задана на опуклій множині M n -вимірного простору, називається **опуклою** на цій множині, якщо

$$F(\alpha X_1 + (1 - \alpha)X_2) \leq \alpha F(X_1) + (1 - \alpha)F(X_2) \quad (6.12)$$

для будь-яких точок $X_1, X_2 \in M$ і будь-якого числа $\alpha \in [0, 1]$.

Якщо в умові (6.12) змінити знак нерівності \leq на \geq , то одержимо визначення **вгнутої** функції. Якщо ж в умові (6.12) нерівність виконується як строга, то функція називається **строго опуклою** (або **строго вгнутою**). На рис. 6.1 зображений графік функції однієї змінної, опуклої на всій числовій прямій.

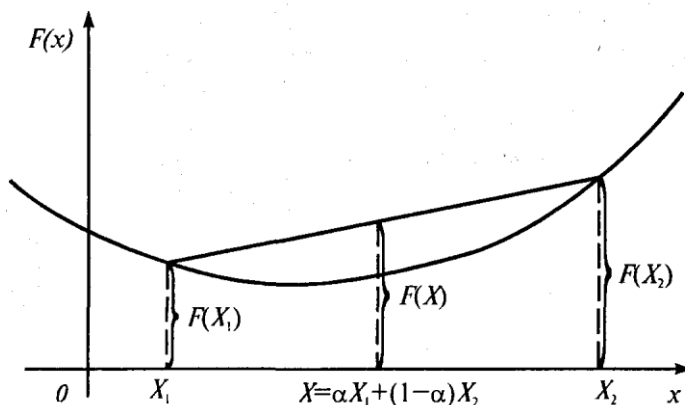


Рисунок 6.1

Алгебраїчні й аналітичні властивості опуклих функцій:

1. Якщо функція $F(X)$ опукла, то функція $-F(X)$ вгнута.
2. Функція $F(X) = C$ і лінійна функція $F(X) = ax + b$ є всюди опуклими й всюди вгнутими.

3. Якщо функції $F_i(X)$, $i = 1, \dots, m$, опуклі, то при будь-яких дійсних числах $\alpha_i \geq 0$ функція $\sum_{i=1}^m \alpha_i F_i(X)$ також є опуклою.

4. Якщо функція $F(X)$ опукла, то для будь-якого числа a область розв'язків нерівності $F(X) < a$ є або опуклою множиною, або порожньою.

Відзначимо наслідок властивості 4 і теореми про перетин опуклих множин.

5. Якщо функції $\varphi_i(X)$ опуклі при всіх невід'ємних значеннях змінних, то область розв'язків системи нерівностей $\varphi_i(X) \leq b_i, i = 1, \dots, m$ є опуклою множиною (якщо вона не порожня).

6. Опукла (вгнута) функція, що визначена на опуклій множині M , є безперервною в кожній внутрішній точці цієї множини.

7. Будь-яка диференційована строго опукла (вгнута) функція має не більше однієї стаціонарної точки (тобто точки, у якій дорівнюють нулю всі частинні похідні). При цьому для опуклої (вгнутої) функції стаціонарна точка завжди є точкою локального й глобального мінімуму (максимуму).

8. Двічі диференційована функція $F(X) = F(x_1, \dots, x_n)$ є опуклою в тому і тільки в тому випадку, коли

$$\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 F(X)}{\partial x_i \partial x_j} \Delta x_i \Delta x_j \geq 0 \quad (6.13)$$

для будь-яких $X \in M$ й $\Delta x_i, \Delta x_j$, що не згортаються в нуль одночасно.

Щоб використати цю умову для визначення опуклості конкретної функції, часто буває корисний критерій Сильвестра: умова (6.13) виконується тоді, коли невід'ємні всі головні мінори Δ_k матриці інших частинних похідних, тобто визначники

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}, a_{ij} = \frac{\partial^2 F(X)}{\partial x_i \partial x_j}, k = 1, \dots, n. \quad (6.14)$$

Якщо всі $\Delta_k > 0$, то нерівність (6.13) виконується як строга, і тоді функція F є строго опуклою.

6.3. Теорема Куна – Таккера

У сучасній теорії пошуку екстремуму опуклих функцій на опуклих множинах класичний метод множників Лагранжа узагальнено. Результатом такого узагальнення є теорема Куна – Таккера. Остання займає центральне місце в теорії нелінійного програмування.

Розглянемо задачу нелінійного програмування, яка може бути розв’язана на основі застосування вищезгаданої теореми.

Знайти максимальне значення функції $Z = f(x_1, x_2, \dots, x_n)$

при обмеженнях

$$g_i(x_1, x_2, \dots, x_n) \geq 0 \quad (i = \overline{1, m}), \quad (6.15)$$

$$x_1, x_2, \dots, x_n \geq 0.$$

При виконанні умови регулярності (існує, принаймні, одна точка X , для якої $g_i(X) > 0$ для всіх i) має місце наступна теорема.

Теорема Куна – Таккера

Вектор $X^{(0)} \geq 0$ тоді й тільки тоді є оптимальним розв’язком задачі (6.15), коли існує такий вектор $\lambda^{(0)} \geq 0$, що для всіх $X \geq 0$ і $\lambda \geq 0$ виконується умова

$$L(X, \lambda^{(0)}) \leq L(X^{(0)}, \lambda^{(0)}) \leq L(X^{(0)}, \lambda). \quad (6.16)$$

У нерівності (6.16) $L(X, \lambda)$ – функція Лагранжа, яка сформована для розглянутої задачі і має вигляд

$$L(X, \lambda) = f(X) + \sum_{i=1}^m \lambda_i g_i(X). \quad (6.17)$$

Точка $(X^{(0)}, \lambda^{(0)})$ називається *сідловою точкою* для функції $L(X, \lambda)$. Оскільки теорема Куна – Таккера встановлює зв’язок між оптимальним планом задачі нелінійного програмування й сідловою точкою функції Лагранжа (6.17), то її ще називають *теоремою про сідлову точку*.

Якщо функції $f(X)$ й $g_i(X)$ диференційовані, то умова (6.16) еквівалентна наступним локальним умовам Куна – Таккера:

$$\left(\frac{\partial L}{\partial x_j}\right)_{X^{(0)}\lambda^{(0)}} \leq 0, \quad x_j^{(0)} \left(\frac{\partial L}{\partial x_j}\right)_{X^{(0)}\lambda^{(0)}} = 0, \quad x_j^{(0)} \geq 0 \quad (j = \overline{1, n}), \quad (6.18)$$

$$\left(\frac{\partial L}{\partial \lambda_j}\right)_{X^{(0)}\lambda^{(0)}} \geq 0, \quad \lambda_j^{(0)} \left(\frac{\partial L}{\partial \lambda_j}\right)_{X^{(0)}\lambda^{(0)}} = 0, \quad \lambda_j^{(0)} \geq 0 \quad (j = \overline{1, m}). \quad (6.19)$$

Вираз $\left(\frac{\partial L}{\partial x_j}\right)_{X^{(0)}\lambda^{(0)}}$ означає, що значення часткової похідної функції Лагранжа $L(X, \lambda)$ знаходиться в точці $(X^{(0)}, \lambda^{(0)})$, де $X^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$, $\lambda^{(0)} = (\lambda_1^{(0)}, \dots, \lambda_m^{(0)})$.

Умови (6.18), (6.19) можуть бути записані у векторній формі в наступному вигляді:

$$\left(\frac{\partial L}{\partial x_j}\right)_{X^{(0)}\lambda^{(0)}} \leq 0, \quad X'^{(0)} \left(\frac{\partial L}{\partial \lambda_j}\right)_{X^{(0)}\lambda^{(0)}} = 0, \quad X^{(0)} \geq 0; \quad (6.20)$$

$$\left(\frac{\partial L}{\partial \lambda_j}\right)_{X^{(0)}\lambda^{(0)}} \geq 0, \quad \lambda'^{(0)} \left(\frac{\partial L}{\partial \lambda_j}\right)_{X^{(0)}\lambda^{(0)}} = 0, \quad \lambda^{(0)} \geq 0. \quad (6.21)$$

Зауваження: при розв'язуванні задач квадратичного програмування, як правило, застосовуються умови (6.20), (6.21).

Контрольні запитання

1. Постановка загальної задачі математичного програмування.
2. Особливості задач нелінійного програмування і методи їх розв'язання.
3. Алгоритм графічного методу розв'язання задач нелінійного програмування.
4. Необхідні та достатні умови екстремуму функції Лагранжа.
5. Алгоритм методу множників Лагранжа.
6. Теорема Куна – Таккера.

7. МОДЕЛІ ДИНАМІЧНОГО ПРОГРАМУВАННЯ

7.1. Постановка задачі динамічного програмування

Динамічне програмування (ДП) – метод оптимізації, пристосований до операцій, у яких процес прийняття рішення може бути розбитий на етапи (кроки). Такі операції називаються *багатокроковими*. Початок розвитку ДП належить до 50-х років XX ст. Воно пов'язане з ім'ям американського математика Р. Беллмана.

Наведемо загальну постановку задачі ДП. Розглядається керований процес. У результаті керування система (об'єкт керування) S переводиться з початкового стану s_0 у стан \tilde{s} . Припустимо, що керування можна розбити на n кроків, тобто рішення приймається послідовно на кожному кроці, а керування, що переводить систему S з початкового стану в кінцеве, являє собою сукупність n пок-рокових керувань.

Позначимо через X_k керування на k -му кроці ($k = 1, 2, \dots, n$). Змінні X_k задовольняють деяким обмеженням й у цьому змісті називаються *припустимими* (X_k може бути числом, точкою в n -вимірному просторі, якісною ознакою).

Нехай $X (X_1, X_2, \dots, X_n)$ — керування, що переводить систему S зі стану s_0 у стан \tilde{s} . Позначимо через s_k стан системи після k -го кроку керування. Одержимо послідовність станів $s_0, s_1, \dots, s_{k-1}, s_k, \dots, s_{n-1}, s_n = \tilde{s}$, що зобразимо кружками (рис. 7.1).

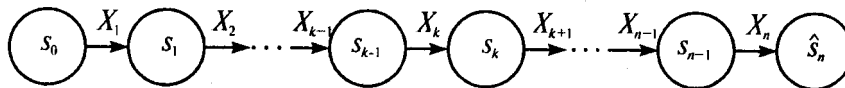


Рисунок 7.1

Показник ефективності розглянутої керованої операції – цільова функція – залежить від початкового стану й керування:

$$Z = F(s_0, X). \quad (7.1)$$

Зробимо кілька припущень.

1. Стан s_k системи наприкінці k -го кроку залежить тільки від попереднього стану s_{k-1} і керування на k -му кроці X_k (і не залежить від попередніх станів і керувань). Ця вимога називається "відсутністю післядії". Сформульоване положення записується у вигляді рівнянь

$$s_k = \varphi_k(s_{k-1}, X_k), k = \overline{1, n}, \quad (7.2)$$

які називаються *рівняннями станів*.

2. Цільова функція (7.1) є адитивною від показника ефективності кожного кроку. Позначимо показник ефективності k -го кроку через

$$Z_k = f_k(s_{k-1}, X_k), k = \overline{1, n}, \quad (7.3)$$

тоді

$$Z = \sum_{k=1}^n f_k(s_{k-1}, X_k). \quad (7.4)$$

Задача покрокової оптимізації (задача ДП) формулюється так: *визначити таке припустиме керування X , що переводить систему S зі стану s_0 у стан \tilde{s} , при якому цільова функція (7.4) набуває найбільшого (найменшого) значення.*

Виділимо особливості моделі ДП:

1. Задача оптимізації інтерпретується як n -кроковий процес керування.
2. Цільова функція дорівнює сумі цільових функцій кожного кроку.
3. Вибір керування на k -му кроці залежить тільки від стану системи до цього кроку, не впливає на попередні кроки (немає зворотного зв'язку).
4. Стан s_k після k -го кроку керування залежить тільки від попереднього стану s_{k-1} і керування X_k (відсутність післядії).
5. На кожному кроці керування X_k залежить від кінцевого числа керуючих змінних, а стан s_k – від кінцевого числа параметрів.

7.2. Принцип оптимальності й рівняння Беллмана

Принцип оптимальності вперше був сформульований Р. Беллманом в 1953 р. Який би не був стан s системи в результаті будь-якого числа кроків, на найближчому кроці потрібно вибирати керування так, щоб воно в сукупності з оптимальним керуванням на всіх наступних кроках приводило до оптимального виграшу на всіх кроках, що залишилися, включаючи даний. Беллманом чітко були сформульовані й умови, за яких принцип правильний. Основна вимога – процес керування повинен бути без зворотного зв'язку, тобто керування на даному кроці не повинно впливати на попередні кроки.

Рівняння Беллмана. Замість вихідної задачі ДП із фіксованим числом кроків n і початковим станом s_0 розглянемо послідовність задач, думаючи послідовно $n = 1, 2, \dots$ при різних s – однокрокову, двокрокову й т.д., – використовуючи принцип оптимальності.

Уведемо ряд нових позначень.

Розглянемо n -й крок: s_{n-1} – стан системи до початку n -го кроку, $s_n = \hat{s}$ – кінцевий стан, X_n – керування на n -му кроці, а $f_n(s_{n-1}, X_n)$ – цільова функція (виграш) n -го кроку. Відповідно до принципу оптимальності X_n потрібно вибирати так, щоб для будь-яких станів s_{n-1} одержати максимум цільової функції на цьому кроці.

Позначимо через $Z_n^*(s_{n-1})$ максимум цільової функції – показника ефективності n -го кроку за умови, що до початку останнього кроку система S була в довільному стані s_{n-1} , а на останньому кроці керування було оптимальним. $Z_n^*(s_{n-1})$ називається *умовним максимумом цільової функції* на n -му кроці. Очевидно, що

$$Z_n^*(s_{n-1}) = \max_{\{X_n\}} f_n(s_{n-1}, X_n). \quad (7.5)$$

Максимізація ведеться за всіма припустимими керуваннями X_n .

Рішення X_n , при якому досягається $Z_n^*(s_{n-1})$ також залежить від s_{n-1} і називається *умовним оптимальним керуванням* на n -му кроці. Воно позначається через $X_n^*(s_{n-1})$. Розв'язавши одновимірну задачу локальної оптимізації за рівнянням (7.5), знайдемо для всіх можливих станів s_{n-1} дві функції: $Z_n^*(s_{n-1})$ і $X_n^*(s_{n-1})$.

Розглянемо тепер двокрокову задачу: приєднаємо до n -го кроку $(n-1)$ -й (рис. 7.2).

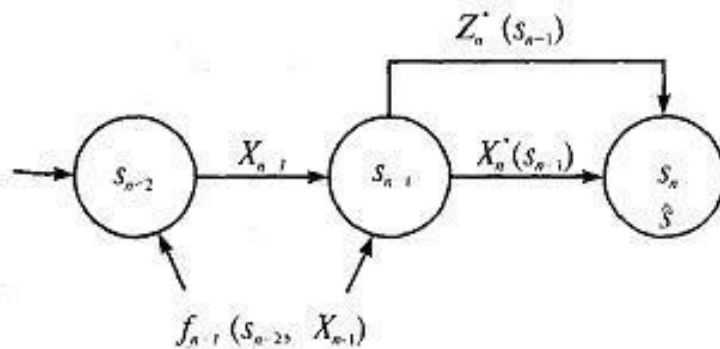


Рисунок 7.2

Умовний максимум цільової функції на $(n-1)$ -му кроці має вигляд

$$Z_{n-1}^*(s_{n-2}) = \max_{\{X_{n-1}\}} \{f_{n-1}(s_{n-2}, X_{n-1}) + Z_n^*(s_{n-1})\}. \quad (7.6)$$

Умовний максимум цільової функції на k -му кроці визначається як

$$Z_k^*(s_{k-1}) = \max_{\{X_k\}} \{f_k(s_{k-1}, X_k) + Z_{k+1}^*(s_k)\}, \quad k = n-1-2, \dots, 2, 1. \quad (7.7)$$

Рівняння (7.7) називають рівняннями Беллмана. Процес розв'язування рівнянь (7.5) і (7.7) називається **умовною оптимізацією**.

Контрольні запитання

1. Постановка задачі динамічного програмування.
2. Особливості моделі динамічного програмування.
3. Принцип оптимальності Беллмана.
4. Рівняння Беллмана.

8. ЕЛЕМЕНТИ ТЕОРІЇ ІГОР

8.1. Предмет і задачі теорії ігор. Основні поняття

Однією з основних проблем дослідження операцій є проблема прийняття рішень в умовах невизначеності. Останні належать до умов виконання операцій, дій сторін, що беруть участь в операції, цілей операції.

Отже, основна задача теорії ігор – обґрунтування розв'язку в умовах невизначеності.

При розв'язанні ряду практичних задач дослідження операцій доводиться аналізувати ситуації, у яких зіштовхуються інтереси двох і більше сторін. Такі ситуації називаються *конфліктними*. Необхідність аналізу таких ситуацій привела до створення математичної теорії конфліктних операцій, яка називається **теорією ігор**.

Математична модель конфліктної ситуації називається грою, сторони, що беруть участь у конфлікті, – *гравцями*, а результат конфлікту – *виграшем*. Для кожної формалізованої гри вводяться *правила*, тобто система умов, що визначає: 1) варіанти дій гравців; 2) обсяг інформації кожного гравця про поведінку партнерів; 3) виграш, до якого приводить кожна сукупність дій. Як правило, виграш (або програш) може бути заданий кількісно, наприклад, можна оцінити програш нулем, виграш – одиницею, а нічию $-1/2$.

Гра називається *парною*, якщо в ній беруть участь два гравці, і *множинною*, якщо число гравців більше двох. Ми будемо розглядати тільки парні ігри. У них беруть участь два гравці A і B , інтереси яких протилежні, а під грою будемо розуміти ряд дій з боку A і B .

Гра називається *грою з нульовою сумою*, або *антагоністичною*, якщо виграш одного із гравців дорівнює програшу іншого, тобто для повного завдання гри досить указати величину одного з них. Якщо позначити a – виграш одного із гравців, b – виграш іншого, то для гри з нульовою сумою $b = -a$, тому досить розглядати, наприклад a .

Вибір і здійснення одного з передбачених правилами дій називається *ходом* гравця. Ходи можуть бути особистими й випадковими. *Особистий хід* – це свідомий вибір гравцем однієї з можливих дій (наприклад, хід у шаховій грі). *Випадковий хід* – це випадково обрана дія (наприклад, вибір карти з перетасованої колоди). Надалі ми будемо розглядати тільки особисті ходи гравців.

Стратегією гравця називається сукупність правил, що визначають вибір його дії при кожному особистому ході залежно від сформованої ситуації. Гра називається *кінцевою*, якщо в кожного гравця є кінцеве число стратегій, і *нескінченною* – у протилежному випадку.

Змішана стратегія – це стратегія, в якій окремі «чисті» стратегії чергуються випадковим чином з деякими ймовірностями.

Для того щоб *розв'язати* гру, або знайти рішення гри, треба для кожного гравця вибрати стратегію, що задовольняє умову *оптимальності*, тобто один із гравців повинен одержувати *максимальний виграш*, коли другий дотримується своєї стратегії. У той же час другий гравець повинен мати *мінімальний програш*, якщо перший дотримується своєї стратегії. Такі *стратегії* називаються *оптимальними*. Оптимальні стратегії повинні також задовольняти умову *стійкості*, тобто кожному із гравців повинно бути не вигідним відмовитися від своєї стратегії в цій грі.

Якщо гра повторюється досить багато разів, то гравців може цікавити не виграш і програш у кожній конкретній партії, а *середній виграш (програш)* у всіх партіях.

Отже, *метою теорії ігор* є визначення *оптимальної стратегії* кожного гравця, а її *предметом* – гра.

8.2. Платіжна матриця. Нижня й верхня ціна гри

Розглянемо парну скінченну гру. Нехай гравець A має m особистих стратегій, які позначимо A_1, A_2, \dots, A_m . Нехай у гравця B є n особистих стратегій, позначимо їх B_1, B_2, \dots, B_n . Говорять, що гра має розмірність $m \times n$. У результаті вибору гравцями будь-якої пари стратегій

$$A_i \text{ і } B_j \quad (i = 1, 2, \dots, m, j = 1, 2, \dots, n)$$

однозначно визначається результат гри, тобто виграш a_{ij} гравця A (позитивний або негативний) і програш $(-a_{ij})$ гравця B . Припустимо, що значення a_{ij} відомі для будь-якої пари стратегій (A_i, B_j) . Матриця $P = (a_{ij})$, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$, елементами якої є виграші, що відповідають стратегіям A_i й B_j , називається *платіжною матрицею* або *матрицею гри*. Загальний вигляд такої матриці поданий у табл. 8.1.

Таблиця 8.1

A_j	B_i			
	B_1	B_2	\dots	B_n
A_1	a_{11}	a_{12}	\dots	a_{1n}
A_2	a_{21}	a_{22}	\dots	a_{2n}
\dots	\dots	\dots	\dots	\dots
A_m	a_{m1}	a_{m2}	\dots	a_{mn}

Рядки цієї таблиці відповідають стратегіям гравця A , а стовпці — стратегіям гравця B .

Розглянемо гру $m \times n$ з матрицею $P = (a_{ij})$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, і визначимо найкращу серед стратегій A_1, A_2, \dots, A_m . Вибираючи стратегію A_i , гравець A повинен розраховувати, що гравець B відповість на неї тією зі стратегій B_j , для якої виграш для гравця A мінімальний (гравець B прагне "нашкодити" гравцеві A).

Позначимо через α_i найменший виграш гравця A при виборі ним стратегії A_i для всіх можливих стратегій гравця B (найменше число в i -тому рядку платіжної матриці), тобто

$$\min_{j=1, \dots, n} a_{ij} = \alpha_i. \quad (8.1)$$

Серед усіх чисел α_i ($i = 1, 2, \dots, m$) виберемо найбільше $\alpha = \max_{i=1, \dots, m} \alpha_i$.

Назвемо α нижньою ціною гри або максимальним виграшем (максимумом). Це гарантований виграш гравця A при будь-якій стратегії гравця B . Отже,

$$\alpha = \max_i \min_j a_{ij}. \quad (8.2)$$

Стратегія, що відповідає максимуму, називається максимінною стратегією. Гравець B зацікавлений у тому, щоб зменшити виграш гравця A , вибираючи стратегію B_j , він ураховує максимально можливий при цьому виграш для A . Позначимо

$$\beta_j = \max_i a_{ij}. \quad (8.3)$$

Серед усіх чисел β_j виберемо найменше $\beta = \min_{j=1, \dots, n} \beta_j$ й назвемо β верхньою ціною гри або мінімаксом (мінімаксом). Це гарантований програш гравця B . Отже,

$$\beta = \min_j \max_i a_{ij}. \quad (8.4)$$

Стратегія, що відповідає мінімаксу, називається *мінімаксною стратегією*.

Фактично виграш гравця A при розумних діях гравців обмежений нижньою й верхньою цінами гри. Якщо вираження рівні, тобто

$$\alpha = \beta \text{ або } \max_i \min_j a_{ij} = \min_j \max_i a_{ij} = v, \quad (8.5)$$

то гра називається *грою із сідловою точкою*, а число v – *ціною гри*.

Елемент $a_{i_0 j_0} = v$ у матриці гри є одночасно мінімальним у рядку i_0 і максимальним у стовпці j_0 та називається *сідловою точкою*.

8.3. Методи розв'язування скінченної гри без сідлової точки

Серед скінченних ігор, що мають практичне значення, ігри із сідловою точкою зустрічаються рідко. Більше типовим є випадок, коли $\alpha \neq \beta$. Такі ігри називаються іграми без сідлової точки.

Розглянемо скінченну гру $m \times n$, платіжна матриця якої подана в таблиці 8.1. У таких іграх, мабуть, $\alpha < \beta$. Якщо кожному гравцеві дати можливість вибору однієї чистої стратегії, то цей вибір повинен визначатися принципом мінімакса. При цьому гравець A гарантує собі виграш рівний α , а гравець B – програш β . Для кожного гравця природним є питання збільшення виграшу (зменшення програшу). Пошуки такого рішення полягають у тому, що гравці застосовують не одну, а кілька стратегій. Вибір стратегій здійснюється випадковим чином. Тобто гравці вибирають змішані стратегії.

У гри, матриця якої має розмірність $m \times n$, стратегії гравця A задаються наборами ймовірностей $X = (x_1, x_2, \dots, x_m)$, з якими гравець застосовує свої початкові чисті стратегії. Ці набори можна розглядати як m -вимірні вектори, для компонентів яких виконуються умови

$$\sum_{i=1}^m x_i = 1, x_i \geq 0 \quad (i = \overline{1, m}). \quad (8.6)$$

Аналогічно стратегії гравця B задаються наборами ймовірностей $Y = (y_1, y_2, \dots, y_n)$, з якими гравець застосовує свої початкові чисті стратегії. Ці набори можна розглядати як n -вимірні вектори, для компонентів яких виконуються умови

$$\sum_{j=1}^n y_j = 1, y_j \geq 0 \quad (j = \overline{1, n}) \quad (8.7)$$

Стратегії гравців A і B , для яких імовірності x_i й y_i відрізняються від нуля, називаються *активними*.

Виграш гравця A при використанні змішаних стратегій визначається як математичне очікування виграшу, тобто дорівнює $\sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j$ або (у векторно-матричному записі) XY' .

Має місце так названа **основна теорема теорії ігор**, що полягає в наступному: кожна кінцева гра має, принаймні, одне рішення в області змішаних стратегій.

Застосування оптимальної стратегії в розглянутій задачі дозволяє одержати виграш, що дорівнює ціні гри:

$$\alpha \leq v \leq \beta.$$

Для оптимальних стратегій гравців має місце співвідношення

$$\max_x \min_y XY' = \min_y \max_x XY'.$$

Спрощення ігор

Загалом кажучи, задача розв'язання гри, якщо її матриця не містить сідлової точки, тим складніша, чим більше значення m й n . Тому в теорії матричних ігор розглядаються способи, за допомогою яких рішення одних ігор зводяться до рішення інших, більш простих (наприклад, за допомогою скорочення розмірності матриці). Скоротити розмірність матриці можна, виключивши дублюючі й заздалегідь не вигідні стратегії.

Дублюючими називаються стратегії, яким відповідають однакові значення елементів у платіжній матриці, тобто якщо в платіжній матриці містяться однакові рядки або стовпці.

Якщо всі елементи i -го рядка матриці менше відповідних елементів k -го рядка, то i -а стратегія для гравця A називається *заздалегідь не вигідною*. Або якщо елементи r -го стовпця матриці більше відповідних елементів j -го стовпця, то для гравця B стратегія B_r заздалегідь не вигідна.

8.4. Зведення матричної гри без сідлової точки до задачі лінійного програмування

Розглянемо скінченну гру без сідлових точки, платіжна матриця якої подана в таблиці 8.1. Оскільки платіжна матриця не містить сідлових точки, то розв'язок гри подається в змішаних стратегіях

$$X = (x_1, x_2, \dots, x_m), Y = (y_1, y_2, \dots, y_n).$$

Застосування гравцем A оптимальної стратегії $X^* = (x_1^*, x_2^*, \dots, x_m^*)$ повинне забезпечити йому при будь-яких діях гравця B виграш не менше ціни гри v . Тому при оптимальній стратегії гравця A повинна виконуватися умова

$$\sum_{i=1}^m x_i * a_{ij} \geq v \left(j = \overline{1, n} \right).$$

Аналогічно для гравця B оптимальна стратегія $Y^* = (y_1^*, y_2^*, \dots, y_n^*)$ повинна забезпечити йому при будь-яких діях гравця A програш, що не перевищує величину v . Тому при оптимальній стратегії гравця B повинна виконуватися умова

$$\sum_{j=1}^n y_j * a_{ij} \leq v \left(i = \overline{1, m} \right).$$

Сформульована задача може бути розглянута як задача знаходження оптимальної стратегії гравця A , для якої виконуються наступні обмеження:

[illegible]

Величина v (ціна гри) невідома, однак можна вважати, що $v > 0$.

Перетворимо систему обмежень (8.8), розділивши всі члени нерівностей на v . У результаті одержимо

[illegible]

$$\text{де} \quad t_i = \frac{x_i}{\nu} \left(i = \overline{1, m} \right). \quad (8.10)$$

З умов $\sum_{i=1}^m x_i = 1, x_i \geq 0 (i = \overline{1, m})$ і виразів (8.10) випливає, що

$$t_1 + t_2 + \dots + t_m = \frac{1}{\nu}. \quad (8.11)$$

Розв'язок гри повинен максимізувати значення v , а значить, із урахуванням (8.11), функція $Z = \sum_{i=1}^m t_i$ повинна набути мінімального значення. Таким чином, розглянута задача теорії ігор зводиться до наступної задачі лінійного програмування:

$$Z = \sum_{i=1}^m t_i \rightarrow \min \quad (8.12)$$

при обмеженнях

[illegible]

$$t_i \geq 0 \left(i = \overline{1, m} \right). \quad (8.14)$$

Невід'ємність змінних $t_i \geq 0 (i = \overline{1, m})$ впливає з (8.10). Рішення ЗЛП (8.12) – (8.14) дозволяє знайти оптимальні значення t_i й величину $\frac{1}{v}$. Це у свою чергу дозволяє знайти значення $x_i = vt_i (i = \overline{1, m})$, які визначають оптимальну стратегію $X^* = (x_1^*, x_2^*, \dots, x_m^*)$ гравця A .

За аналогією для знаходження стратегії гравця B треба врахувати умову $\sum_{j=1}^n y_j * a_{ij} \leq v(i = \overline{1, m})$. Її врахування дозволяє звести задачу про пошук стратегії гравця B до наступної ЗЛП:

$$W = \sum_{j=1}^n u_j \rightarrow \max \quad (8.15)$$

при обмеженнях

[illegible]

$$u_j \geq 0 \quad (j = \overline{1, n}). \quad (8.17)$$

У ЗЛП (8.15) – (8.17)

$$u_j = \frac{y_j}{v} \quad (j = \overline{1, n}).$$

Таким чином, розв'язок сформульованої гри зводиться до розв'язання пари двоїстих симетричних ЗЛП. Використовуючи властивість симетричності, можна розв'язати одну з них, а саме ту, для якої потрібно менше обчислень, а розв'язок іншої задачі знайти на основі співвідношень двоїстості.

Контрольні запитання

1. Поняття теорії ігор та її основної задачі, гри, гравців, виграшу та правил.
2. Поняття ходу гравця, його стратегії, рішення гри, умови стійкості оптимальної стратегії, мети теорії ігор.
3. Платіжна матриця. Нижня та верхня ціна гри.
4. Методи розв'язування скінченної гри без сідлової точки.
5. Зведення матричної гри без сідлової точки до задачі лінійного програмування.

9. ПРЕДМЕТ, МЕТОДИ ТА ЗАДАЧІ ЕКОНОМЕТРІЇ. ОСОБЛИВОСТІ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ

Економетрія – це один з базових курсів підготовки економістів. Ця дисципліна ґрунтується на фундаменті знань економічної теорії, матричної алгебри, теорії математичної й загальної статистики.

Економетрія – це один з напрямків економіко-математичних методів аналізу, що полягає в статистичному вимірі (оцінюванні) параметрів математичних виражень, які характеризують деяку економічну концепцію про взаємозв'язок і розвиток об'єкта, явища, що необхідно для одержання конкретних економічних висновків на основі економетричних моделей.

Економетрія – це наука, що досліджує кількісні закономірності й взаємозалежності в економіці за допомогою методів математичної статистики. Основа цих методів – кореляційно-регресійний аналіз.

Методологія економетричного моделювання тісно пов'язана з використанням системного підходу. Соціально-економічна система може бути подана нескінченним числом структурних і функціональних інваріантів, що відображають взаємозв'язки між різними процесами, що відбуваються у цій системі (економічними, соціальними, екологічними, демографічними й т.д.). Опис системи здійснюється за допомогою її якісних і кількісних характеристик, що називаються *параметрами*. Параметри становлять основу мов опису систем, а при формалізації ототожнюються з незалежними змінними математичного опису процесу функціонування систем.

При побудові економетричної моделі реалізується метод моделювання за принципом «чорного ящика», коли дослідникові невідомий механізм процесів, що відбуваються в системі, вивчити який можна за вхідними і вихідними характеристиками системи. Вхідні та вихідні характеристики системи часто ототожнюють із екзогенними й ендогенними змінними або в кореляційно-регресійному аналізі вживають терміни «незалежні (факторні) змінні» або ознаки й «залежні (результативні) змінні» або ознаки. Графічно принцип «чорного ящика» зображений на рис. 9.1.

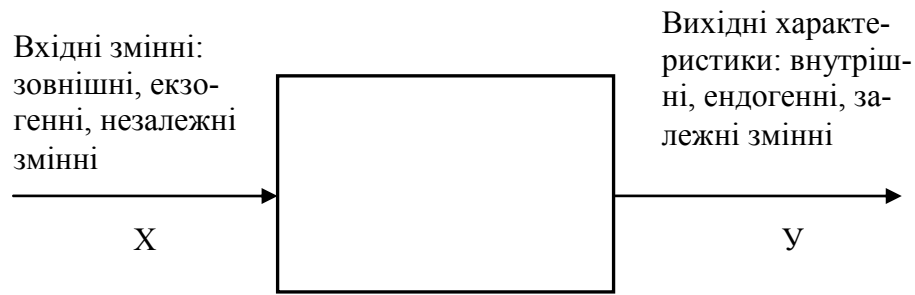


Рисунок 9.1

Досліднику необхідно виділити вхідні й вихідні характеристики та на підставі економетричних методів установити характер причинно-наслідкових зв'язків, що лежать в основі механізму функціонування соціально-економічної системи.

Заключним етапом у процедурі послідовної формалізації опису процесів функціонування соціально-економічних систем є розробка математичних моделей.

9.2. Особливості побудови економетричних моделей

Економетрична модель – це особливий клас економіко-математичних моделей, у яких дослідник вирішує цілий ряд задач:

- вибір форми математичної залежності, що описує поведіння економічного об'єкта на основі системи спостережень;
- оцінка параметрів даної моделі різними методами (метод найменших квадратів, метод максимальної правдоподібності та ін.);
- перевірка статистичної значущості моделі.

Часто *економетрична модель у загальному вигляді* подається як система лінійних рівнянь:

$$BY = AX + E,$$

де B – матриця коефіцієнтів при ендогенних (залежних) змінних;

Y – вектор ендогенних змінних (спостережуване значення залежної змінної);

A – матриця коефіцієнтів при екзогенних (пояснюючих) змінних;

X – вектор екзогенних змінних;

AX – пояснювальна частина, що залежить від значень пояснюючих змінних;

E – вектор випадкових збурювань (помилки, відхилення).

Економетричні моделі включають досить широкий клас різних економіко-математичних моделей. Наведемо одну із **класифікацій економетричних моделей**.

1. За способом математичного подання економетричні моделі можна умовно розділити на прості й складні. *Прості* економетричні моделі подані одним рівнянням, однією залежністю, *складні* – декількома рівняннями, декількома залежностями.

2. За кількістю факторних ознак, що включає модель, прості економетричні моделі можна розділити на однофакторні й багатфакторні. *Однофакторні моделі* містять одну незалежну ознаку, *багатфакторні моделі* – ряд незалежних ознак. Однофакторні й багатфакторні моделі можуть бути подані лінійними та нелінійними функціями.

3. Складні економетричні моделі можуть бути подані трьома видами систем одночасних рівнянь залежно від форми включення до правої частини ендогенних змінних. Звичайно виділяють три типи систем:

- 1) системи, що розв'язані відносно ендогенних змінних;
- 2) рекурсивні системи;
- 3) системи, що не розв'язані відносно ендогенних змінних.

4. Залежно від наявності (відсутності) у моделі фактора часу розрізняють динамічні й статичні моделі. Прикладами динамічних моделей є: *трендові моделі; моделі згладжування часових рядів; моделі декомпозиції часового ряду; авторегресійні моделі й моделі ковзної середньої; лагові моделі та динамічні регресійні моделі.*

Побудова економетричної моделі проводиться у кілька **основних етапів**:

1) якісний аналіз (постановка мети аналізу, визначення сукупності, визначення результативних і факторних ознак, вибір періоду, за який проводиться аналіз, вибір методу аналізу);

2) попередній аналіз сукупності, що моделюється (перевірка однорідності сукупності, виключення аномальних спостережень, уточнення необхідного обсягу ознак, установлення законів розподілу ознак);

3) побудова економетричної моделі (установлення переліку факторів, розрахунок оцінок параметрів рівнянь регресії, перебір конкуруючих варіантів моделі);

4) оцінка адекватності моделі (перевірка статистичної істотності рівняння залежності в цілому та його окремих параметрів; перевірка відповідності формальних властивостей оцінок завданням дослідження);

5) економічна інтерпретація та практичне використання моделі.

Контрольні запитання

1. Предмет економетрії.
2. Параметри системи, екзогенні та ендогенні змінні.
3. Економетрична модель, її загальний вигляд.
4. Класифікація та етапи побудови економетричних моделей.

10. ПАРНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

10.1. Основні поняття

Якщо певному значенню величини X відповідає розподіл деякої величини Y , то між ними існує *статистична (стохастична, імовірнісна) залежність*.

Величина Y називається *залежною змінною, результуючим фактором (ознакою), ендогенною змінною*.

Величина X називається *незалежною змінною, екзогенною змінною*.

Залежність математичного очікування випадкової величини Y від значень змінної X називається *кореляційною залежністю*.

$$M_x(Y) = \varphi(x). \quad (10.1)$$

Рівняння (10.1) називається рівнянням регресії, $\varphi(x)$ – функцією регресії, її графік – *лінією регресії*.

За результатами дослідів можна одержати:

$$\hat{y} = \hat{\varphi}(x, b_0, b_1, \dots, b_p), \quad (10.2)$$

де \hat{y} – умовна (групова) середня змінної Y ; b_0, b_1, \dots, b_p – параметри рівняння регресії.

Рівняння (10.2) називається *вибірковим рівнянням регресії*.

10.1. Лінійна парна регресія

Рівняння лінійної парної регресії має вигляд:

$$\hat{y} = b_0 + b_1 x. \quad (10.3)$$

Відповідно до **методу найменших квадратів (МНК)** невідомі параметри b_0 й b_1 вибираються таким чином, щоб сума квадратів відхилень емпіричних значень y_i від значень \hat{y}_i , що знайдені за рівнянням регресії (10.3), була мінімальною:

$$S = \sum_{i=1}^n (\hat{y}_i - y_i)^2 = \sum_{i=1}^n (b_0 + b_1 x_i - y_i)^2 \rightarrow \min. \quad (10.4)$$

На підставі необхідної умови екстремуму функції двох змінних $S = S(b_0, b_1)$ (10.4) прирівнюємо до нуля її частинні похідні, тобто

$$\begin{cases} \frac{\partial S}{\partial b_0} = 2 \sum_{i=1}^n (b_0 + b_1 x_i - y_i) = 0; \\ \frac{\partial S}{\partial b_1} = 2 \sum_{i=1}^n (b_0 + b_1 x_i - y_i) x_i = 0, \end{cases}$$

звідки після перетворень одержимо систему нормальних рівнянь для визначення параметрів лінійної системи:

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i; \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i. \end{cases} \quad (10.5)$$

Тепер, розділивши обидві частини рівняння (10.5) на n , одержимо систему нормальних рівнянь у вигляді:

$$\begin{cases} b_0 + b_1 \bar{x} = \bar{y}; \\ b_0 \bar{x} + b_1 \bar{x}^2 = \overline{xy}, \end{cases} \quad (10.6)$$

де відповідні середні визначаються за формулами:

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^n x_i}{n}; & \overline{xy} &= \frac{\sum_{i=1}^n x_i y_i}{n}; \\ \bar{y} &= \frac{\sum_{i=1}^n y_i}{n}; & \overline{x^2} &= \frac{\sum_{i=1}^n x_i^2}{n}. \end{aligned}$$

Підставляючи значення $b_0 = \bar{y} - b_1 \bar{x}$ з першого рівняння системи (10.6) у рівняння регресії (10.3) одержимо

$$b_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{K_{xy}}{s_x^2}, \quad (10.7)$$

де b_1 – вибірковий коефіцієнт регресії; K_{xy} – вибірковий кореляційний момент або вибірка кореляція; s_x^2 – вибірка дисперсія змінної X :

$$K_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}, \quad s_x^2 = \overline{x^2} - \bar{x}^2.$$

Якщо значення змінних великі, то розрахунки можуть бути проведені у відхиленнях від середніх величин.

$$\begin{aligned} x'_i &= x_i - \bar{x}; \\ y'_i &= y_i - \bar{y}; \end{aligned} \quad \hat{y}_i = b_1 \cdot x'_i;$$

$$b_1 = \frac{\overline{x'y'}}{x'^2};$$

$$\frac{\hat{y} - \bar{y}}{s_y} = \frac{b_1}{s_y} \cdot \frac{(x - \bar{x})}{s_x} \cdot s_x.$$

Величина $r = \frac{b_1 \cdot s_x}{s_y}$ називається *вибірковим коефіцієнтом кореляції*.

Він показує, наскільки величин s_y зміниться Y , якщо X зміниться на одне s_x .

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{s_x \cdot s_y}. \quad (10.8)$$

Підставивши у вираз (10.8) вихідні дані, одержимо

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \cdot \sqrt{n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2}}. \quad (10.9)$$

Коефіцієнт кореляції має такі властивості:

1) Коефіцієнт кореляції набуває значення на відрізку $[-1;1]$, тобто $-1 \leq r \leq 1$. Чим ближче $|r|$ до 1, тим тісніше кореляційний зв'язок.

2) При $|r| = 1$ кореляційний зв'язок стає функціональним. При цьому всі спостережувані значення лежать на одній лінії.

3) При $r = 0$ кореляційний зв'язок відсутній та лінія регресії паралельна осі x . При $r > 0$ ($b_1 > 0$) кореляційний зв'язок називається *прямим*. При $r < 0$ ($b_1 < 0$) кореляційний зв'язок називається *оберненим*.

Контрольні запитання

1. Алгоритм методу найменших квадратів для парної регресії.
2. Коефіцієнт кореляції, його властивості.

11. ОСНОВНІ ПОЛОЖЕННЯ РЕГРЕСІЙНОГО АНАЛІЗУ. ОЦІНКА ПАРАМЕТРІВ РЕГРЕСІЙНОЇ МОДЕЛІ. ТЕОРЕМА ГАУССА–МАРКОВА. ІНТЕРВАЛЬНА ОЦІНКА ФУНКЦІЇ РЕГРЕСІЇ ТА ЇЇ ПАРАМЕТРІВ. ОЦІНКА ЗНАЧУЩОСТІ РІВНЯННЯ РЕГРЕСІЇ

11.1. Основні положення регресійного аналізу

Як відзначено в розділі 10, розглянута в регресійному аналізі залежність Y від X може бути подана у вигляді модельного рівняння регресії:

$$M_x(Y) = \varphi(X).$$

Окремі виміри величини Y будуть відрізнятися від обчислених значень за рахунок неврахованих факторів і помилок спостереження

$$Y = \varphi(X) + \varepsilon,$$

де ε – випадкова величина, що називається *збурюванням або помилкою*.

Для лінійної моделі ці рівняння мають вигляд:

$$M_x(Y) = \beta_0 + \beta_1 \cdot x. \quad (11.1)$$

Спостережувані значення величини y визначаються за формулою

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon. \quad (11.2)$$

Основні передумови регресійного аналізу:

1. У моделі (11.2) збурювання ε_i (або y_i) є випадковою величиною, а x_i – не випадковою величиною.

2. Математичне очікування збурювання ε_i дорівнює нулю:

$$M(\varepsilon_i) = 0, \quad M(y_i) = \beta_0 + \beta_1 x_i.$$

3. Дисперсія збурювання ε_i (або y_i) є постійною величиною для будь-якого номера i :

$$D(\varepsilon_i) = \sigma^2 \text{ або } D(y_i) = \sigma^2.$$

Ця умова називається умовою *гомоскедастичності* (рівномірності).

4. Збурювання ε_i та ε_j некорельовані:

$$M(\varepsilon_i \cdot \varepsilon_j) = 0, \quad \text{якщо } i \neq j.$$

5. Збурювання ε_i (або y_i) мають нормальні розподіли.

11.2. Оцінка параметрів регресійної моделі

Вибірковою оцінкою рівняння (11.2) є рівняння регресії:

$$\hat{y} = b_0 + b_1 x.$$

Вплив неврахованих факторів і помилок спостережень у моделі (11.2) визначається за допомогою дисперсії збурювань або залишкової дисперсії.

Незмщеною оцінкою залишкової дисперсії є:

$$s^2 = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n - m} = \frac{\sum_{i=1}^n e_i^2}{n - 2}, \quad (11.3)$$

де y_i – спостережуване значення; \hat{y}_i – групова середня, що знайдена за рівнянням регресії; $e_i = \hat{y}_i - y_i$ – вибіркова оцінка збурювання ε_i або залишок регресії; $m = 2$ – число зв'язків (два рівняння для визначення b_0, b_1).

11.3. Теорема Гаусса – Маркова

Теорема Гаусса – Маркова. Якщо регресійна модель (11.2) задовольняє умови 1–4 підрозділу 11.1, то оцінки b_0, b_1 мають найменшу дисперсію в класі всіх незміщених лінійних оцінок.

11.4. Інтервальна оцінка функції регресії та її параметрів

1. Знайдемо довірчий інтервал для умовного математичного очікування $M_x(Y)$ із заданою надійністю.

Припустимо, що γ – це задана надійність. Тоді мають місце наступні рівняння:

$$\hat{y} - \bar{y} = b_1(x - \bar{x}) \text{ – рівняння регресії;}$$

$$\hat{y} = \bar{y} + b_1(x - \bar{x});$$

$$\sigma_{\hat{y}}^2 = \sigma_{\bar{y}}^2 + (x - \bar{x})^2 \cdot \sigma_{b_1}^2 \text{ – рівняння дисперсії.}$$

Можна показати, що

$$\sigma_{\bar{y}}^2 = \frac{\sigma^2}{n}, \quad \sigma_{b_1}^2 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Тоді оцінкою дисперсії $\sigma_{\hat{y}}^2$ є:

$$s_{\hat{y}}^2 = s^2 \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right),$$

де s^2 – вибіркова оцінка σ^2 .

Можна показати, що статистика $t = \frac{\hat{y} - M_x(Y)}{s_{\hat{y}}}$ розподілена за t -розподілом Стюдента з k ступенями волі. За таблицею цього розподілу для рівня значущості α та числа ступенів волі k знаходимо $t_{1-\alpha; k}$. Тоді довірчий інтервал для $M_x(Y)$ має вигляд:

$$\hat{y}_0 - t_{1-\alpha; k} \cdot s_{\hat{y}_0} \leq M_x(Y) \leq \hat{y}_0 + t_{1-\alpha; k} \cdot s_{\hat{y}_0},$$

$$\text{де } s_{\hat{y}_0}^2 = s^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right).$$

Прогнозовані значення залежної змінної Y залежать від відхилення незалежного фактора осі \bar{X} . Чим більше це відхилення (останнє відхилення), тим більше помилка прогнозу.

2. Знайдемо довірчий інтервал для індивідуального значення y_0^* :

$$s_{y_0^*}^2 = s_{\hat{y}}^2 + s^2;$$

$$s_{y_0^*}^2 = s^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right).$$

Довірчий інтервал для y_0^* на рівні значущості α та при числі ступенів волі k має вигляд:

$$\hat{y}_0 - t_{1-\alpha; k} \cdot s_{y_0^*} \leq y_0^* \leq \hat{y}_0 + t_{1-\alpha; k} \cdot s_{y_0^*}.$$

3. Знайдемо довірчий інтервал для параметра β_1 .

Можна показати, що статистика $t = \frac{b_1 - \beta_1}{s} \cdot \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}$ має

t -розподіл Стюдента. Тоді довірчий інтервал для параметра β_1 на рівні значущості α та при числі ступенів волі k має вигляд:

$$b_1 - t_{1-\alpha/2; k} \cdot \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \leq \beta_1 \leq b_1 + t_{1-\alpha/2; k} \cdot \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}.$$

4. Знайдемо довірчий інтервал для дисперсії збурювань σ^2 .

Можна показати, що статистика $\frac{ns^2}{\sigma}$ має χ^2 -розподіл зі $k = n - 2$ ступенями волі. Тоді довірчий інтервал для σ^2 на рівні значущості α має вигляд:

$$\frac{ns^2}{\chi_{\alpha/2; n-2}^2} \leq \sigma^2 \leq \frac{ns^2}{\chi_{1-\alpha/2; n-2}^2}.$$

11.5. Оцінка значущості рівняння регресії. Коефіцієнт детермінації

Перевірити значущість рівняння регресії – значить установити, чи відповідає математична модель експериментальним даним і чи досить включених у рівняння пояснюючих змінних (однієї або декількох) для опису результатуючих факторів.

Перевірка значущості рівняння регресії може бути проведена способом оцінки значущості коефіцієнта регресії b_1 , що має t -розподіл з k ступенями волі. Рівняння парної регресії (або коефіцієнта регресії b_1) значуще на рівні α , якщо спостережуване значення статистики $t = \frac{b_1}{s} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}$ за абсолютною величиною більше табличного значення $t_{1-\alpha/2; k}$.

Обчислимо суму квадратів відхилень.

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n [(\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)]^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2 + 2(\hat{y}_i - \bar{y})(y_i - \hat{y}_i).$$

Можна показати, що третій доданок дорівнює нулю, тоді це рівняння запишемо так:

$$Q = Q_R + Q_e,$$

де Q – загальна сума квадратів відхилень залежної змінної від середньої; Q_R – сума квадратів, що обумовлена регресією; Q_e – сума квадратів, що обумовлена неврахованими факторами.

Ефективною оцінкою адекватності регресійної моделі є *коефіцієнт детермінації*, який обчислюється за допомогою наступної формули:

$$R^2 = \frac{Q_R}{Q} = 1 - \frac{Q_e}{Q}.$$

Коефіцієнт детермінації показує, яка частина зміни залежної перемінної обумовлена зміною незалежного фактора.

$$0 \leq R^2 \leq 1.$$

Чим ближче R^2 до 1, тим краще рівняння регресії наближає (апроксимує) експериментальні дані. У випадку парної регресії $R^2 = r^2$, де r – коефіцієнт кореляції.

11.6. Рангова кореляція. Коефіцієнт Спірмена

Якщо між двома змінними, що виражені своїми рангами, існує кореляційний зв'язок, то він називається *ранговою кореляцією*. Тісноту такого зв'язку визначає коефіцієнт рангової кореляції Спірмена:

$$\rho = 1 - \frac{6 \sum_{i=1}^n (r_i - s_i)^2}{n^3 - n},$$

де r_i, s_i – ранги i -го об'єкта за двома показниками; n – число спостережень.

Контрольні запитання

1. Основні передумови регресійного аналізу.
2. Залишкова дисперсія, теорема Гаусса – Маркова.
3. Інтервальна оцінка функції регресії та її параметрів.
4. Оцінка значущості рівняння регресії. Коефіцієнт детермінації.
5. Рангова кореляція. Коефіцієнт Спірмена.

12. МНОЖИННИЙ РЕГРЕСІЙНИЙ АНАЛІЗ.

12.1. Елементи лінійної алгебри

Матрицею розміру $(m \times n)$ називається таблиця чисел, що містить m рядків й n стовпців:

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Дії з матрицями:

1. Множення матриць.

$$C_{m \times l} = A_{m \times n} \cdot B_{n \times l}.$$

1.1. Множити можна тільки погоджені матриці (повинне $n = p$).

1.2. Розмір матриці-добутку буде дорівнювати $m \times l$.

$$1.3. \quad C_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj}.$$

2. Знаходження оберненої матриці.

2.1. Обернена матриця існує тільки для невиродженої квадратної матриці (якщо $\Delta A = 0$ – то вироджена, якщо $\Delta A \neq 0$ – не вироджена, $m = n$ – квадратна матриця).

2.2. Алгоритм знаходження оберненої матриці

$$A^{-1} \cdot A = E,$$

де E – одинична матриця до матриці A .

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

1) Знайти визначник матриці ΔA .

2) Знайти транспоновану матрицю A' .

- 3) Знайти *приєднану матрицю* A^S , її елементами є алгебраїчні доповнення відповідних елементів матриці A' .
- 4) $A^{-1} = \frac{1}{\Delta A} \cdot A^S$.

12.2. Класична нормальна лінійна модель

Нехай Y – залежна змінна. Тоді змінні фактори, від яких залежить величина Y будемо позначати: X_1, X_2, \dots, X_p .

Будемо припускати, що зв'язок між цими змінними лінійний, тоді *модель множинної регресії* має вигляд:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad (12.1)$$

де $i = \overline{1, n}$; y_i – залежна змінна; x_i – i -е спостереження відповідних факторів.

Ця модель задовольняє основні передумови регресійного аналізу (розділ 11).

Уведемо наступні позначення:

1) $Y = (y_1, y_2, \dots, y_n)'$ називається матрицею-стовпцем (або вектором) залежної змінної Y .

$$2) X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} - \text{називається матрицею значень факторів}$$

або матрицею плану розміру $n \times (p+1)$.

3) $\beta = (\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_p)'$ – матриця-стовпець (вектор) параметрів.

4) $\varepsilon = (\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_p)'$ – матриця-стовпець (вектор) збурювань.

Тоді система рівнянь (12.1) у матричній формі має вигляд:

$$Y = X\beta + \varepsilon. \quad (12.2)$$

Вибірковою оцінкою цієї моделі є

$$Y = Xb + e, \quad (12.3)$$

де $b = (b_0 \ b_1 \ \dots \ b_p)'$, $e = (e_1 \ e_2 \ \dots \ e_n)'$.

Для знаходження вектора оцінок параметрів b застосуємо МНК.

Сума квадратів відхилень дорівнює:

$$S = \sum_{i=1}^n (\hat{y}_i - y_i)^2 = \sum_{i=1}^n e_i^2.$$

$$\sum_{i=1}^n e_i^2 = (e_1 \ e_2 \ \dots \ e_n) \cdot (e_1 \ e_2 \ \dots \ e_n)'$$

$$\sum_{i=1}^n e_i^2 = (Y - Xb) \cdot (Y - Xb)' \rightarrow \min.$$

Остаточно маємо: $S = (Y - Xb) \cdot (Y - Xb)'$.

Можна показати, що вектор частинних похідних дорівнює

$$\frac{\partial S}{\partial b} = -2X'Y + 2X'Xb.$$

Дорівнявши похідну до нуля, одержуємо систему нормальних рівнянь у матричній формі для знаходження вектора b :

$$X'Xb = X'Y. \quad (12.4)$$

Розв'язком цього матричного рівняння є вектор

$$b = (X'X)^{-1} X'Y. \quad (12.5)$$

Рівняння (12.4) правильне при виконанні умов:

- 1) вектор збурювань ε – випадковий вектор, X – не випадкова матриця.
- 2) $M(\varepsilon) = 0_n$.
- 3, 4) $\Sigma_\varepsilon = M(\varepsilon\varepsilon') = \sigma^2 E_n$.
- 5) ε – нормально розподілений випадковий вектор.

б) Щоб існувала обернена матриця $(X'X)^{-1}$, необхідна умова: $\Delta(X'X)^{-1} \neq 0$.

У цьому випадку ранг $r(X) = p + 1 < n$,

де n – число спостережень повинне перевищувати ранг для надійності одержуваних результатів.

Модель, що задовольняє ці умови, називається *класичною нормальною лінійною моделлю множинної регресії*. Якщо не виконується умова 5, то в цій назві викидається слово «нормальна».

При відомому векторі b рівняння множинної регресії має вигляд:

$$\hat{y}_0 = X'_0 b,$$

де $X'_0 = (1 \quad x_{10} \quad x_{20} \quad \dots \quad x_{p0})$.

Для порівняння впливу на залежну змінну різних факторів використовують стандартизовані коефіцієнти регресії b'_j й еластичності E'_j ($j = \overline{1, p}$).

$$b'_j = b_j \frac{S_{x_j}}{S_y};$$

$$E'_j = b_j \frac{\bar{x}_j}{\bar{y}}.$$

$$S_{x_j} = \bar{x}_j^2 - (\bar{x})^2; \quad S_y = \bar{y}^2 - (\bar{y})^2.$$

Перетворимо вектор оцінок (12.5):

$$b = (X'X)^{-1} X'Y.$$

$$Y = X\beta + \varepsilon.$$

$$b = (X'X)^{-1} X'(X\beta + \varepsilon) = (X'X)^{-1} X'X\beta + (X'X)^{-1} X'\varepsilon =$$

$$= E\beta + (X'X)^{-1} X'\varepsilon = \beta + (X'X)^{-1} X'\varepsilon.$$

Оцінка параметра b містить випадкові помилки.

$$M(b) = M(\beta) + M[(X'X)^{-1} X'\varepsilon].$$

$$M(b) = \beta + (X'X)^{-1} X' M(\varepsilon).$$

$$M(\varepsilon) = 0 \text{ — за умовами математичної моделі.}$$

$$M(b) = \beta.$$

12.3. Коваріаційна матриця

Коваріаційна матриця визначає розсіювання оцінок рівняння регресії.

У загальному вигляді вона записується так:

$$\Sigma_b = \begin{pmatrix} \sigma_{00} & \sigma_{01} & \dots & \sigma_{0p} \\ \sigma_{10} & \sigma_{11} & \dots & \sigma_{1p} \\ \dots & \dots & \dots & \dots \\ \sigma_{p0} & \sigma_{p1} & \dots & \sigma_{pp} \end{pmatrix}, \quad (12.6)$$

де $\sigma_{ij} = M[(b_i - M(b_i))(b_j - M(b_j))]$ — коваріації (або кореляційні моменти двох змінних b_i, b_j).

Коваріація показує ступінь розсіювання цих змінних і їхній взаємозв'язок.

$$\sigma_{ij} = M[(b_i - \beta_i)(b_j - \beta_j)].$$

На головній діагоналі елементи матриці (5.6):

$$\sigma_{jj} = \sigma_{b_j} = M[(b_j - \beta_j)^2]$$

Тоді матрицю (12.6) можна подати як

$$\Sigma_b = M[(b - \beta)(b - \beta)'].$$

$$b = \beta + (X'X)^{-1} X' \varepsilon.$$

$$\begin{aligned} \Sigma_b &= M\left[\left((X'X)^{-1} X' \varepsilon\right)\left((X'X)^{-1} X' \varepsilon\right)'\right] = M\left[(X'X)^{-1} X' \varepsilon \varepsilon' X (X'X)^{-1}\right] = \\ &= (X'X)^{-1} X' M(\varepsilon \varepsilon') X (X'X)^{-1}. \end{aligned}$$

Матриця $M(\varepsilon \varepsilon')$ являє собою коваріаційну матрицю вектора збурювань ε .

$$M(\varepsilon\varepsilon') = \begin{pmatrix} M(\varepsilon_1^2) & M(\varepsilon_1\varepsilon_2) & \dots & M(\varepsilon_1\varepsilon_n) \\ M(\varepsilon_2\varepsilon_1) & M(\varepsilon_2^2) & \dots & M(\varepsilon_2\varepsilon_n) \\ \dots & \dots & \dots & \dots \\ M(\varepsilon_n\varepsilon_1) & M(\varepsilon_n\varepsilon_2) & \dots & M(\varepsilon_n^2) \end{pmatrix}. \quad (12.7)$$

За умовами 3, 4:

$$M(\varepsilon\varepsilon') = \sigma^2 E_n,$$

E_n – одинична матриця n -го порядку.

$$\Sigma_b = (X'X)^{-1} X' \sigma^2 E_n X (X'X)^{-1} = \sigma^2 (X'X)^{-1}.$$

Матриця $(X'X)^{-1}$ визначає не тільки вектор b , але й дисперсії його компонентів.

Контрольні запитання

1. Алгоритм знаходження оберненої матриці.
2. Алгоритм методу найменших квадратів для множинної регресії.
3. Умови існування класичної нормальної лінійної моделі множинної регресії.
4. Коваріаційна матриця.

13. ДОВІРЧІ ІНТЕРВАЛИ ДЛЯ КОЕФІЦІЄНТІВ І ФУНКЦІЇ РЕГРЕСІЇ. ОЦІНКА ЗНАЧУЩОСТІ РІВНЯННЯ МНОЖИННОЇ РЕГРЕСІЇ. МУЛЬТИКОЛІНЕАРНІСТЬ. ЛІНІЙНІ РЕГРЕСІЙНІ МОДЕЛІ ЗІ ЗМІННОЮ СТРУКТУРОЮ. ФІКТИВНІ ЗМІННІ

13.1. Довірчі інтервали для коефіцієнтів і функції регресії

Оцінкою $s_{b_j}^2$ дисперсії $\sigma_{b_j}^2$ є величина:

$$s_{b_j}^2 = s^2 [(X'X)^{-1}]_{jj},$$

де $s^2 = \frac{\sum_{i=1}^n e_i^2}{n-p-1}$ – незміщена оцінка σ^2 ; $[(X'X)^{-1}]_{jj}$ – діагональний елемент оберненої матриці.

Коефіцієнт b_j є значущим на рівні α , якщо

$$|t| = \frac{|b_j|}{s_{b_j}} > t_{1-\alpha/2; k}.$$

Тоді довірчий інтервал параметра β_j має вигляд

$$b_j - t_{1-\alpha/2; k} \cdot s_{b_j} \leq \beta_j \leq b_j + t_{1-\alpha/2; k} \cdot s_{b_j}.$$

1. Знайдемо довірчий інтервал для математичного очікування $M_{X_0}(Y)$ за умови, що вектор X набув значення:

$$X'_0 = (1 \quad x_{10} \quad x_{20} \quad \dots \quad x_{p0}).$$

Знаходимо з рівняння регресії групову середню величину, що ділена на \hat{y}_0 .

$$s_{\hat{y}_0} = s \sqrt{X'_0 (X'X)^{-1} X_0}.$$

Тоді довірчий інтервал має вигляд

$$\hat{y}_0 - t_{1-\alpha/2; k} \cdot s_{\hat{y}_0} \leq M_X(Y) \leq \hat{y}_0 + t_{1-\alpha/2; k} \cdot s_{\hat{y}_0}.$$

2. Знайдемо довірчий інтервал для індивідуального значення y_0^* :

$$s_{y_0^*}^2 = s_{\hat{y}_0}^2 + s^2.$$

$$\hat{y}_0 - t_{1-\alpha/2; k} \cdot s_{y_0^*} \leq y_0^* \leq \hat{y}_0 + t_{1-\alpha/2; k} \cdot s_{y_0^*}.$$

3. Знайдемо довірчий інтервал для дисперсії σ^2 :

$$\frac{ns^2}{\chi_{\frac{\alpha}{2}; k}} \leq \sigma^2 \leq \frac{ns^2}{\chi_{1-\frac{\alpha}{2}; k}}.$$

13.2. Оцінка значущості рівняння множинної регресії

Обчислимо коефіцієнт детермінації множинної регресії R^2 .

Сума загального числа квадратів відхилень дорівнює:

$$Q = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}.$$

$$\sum_{i=1}^n y_i^2 = Y'Y.$$

$$Q = Y'Y - n\bar{y}^2.$$

Суми квадратів відхилень, що обумовлені неврахованими факторами Q_e й регресією Q_R , дорівнюють:

$$Q_e = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 = e'e = (Y - Xb)'(Y - Xb) = |X'Xb = X'Y| = Y'Y - b'X'Y.$$

$$Q_R = Q - Q_e = b'X'Y - n(\bar{y})^2.$$

Тоді коефіцієнт детермінації розраховується за формулою

$$R^2 = Q_R / Q = \frac{b'X'Y - n(\bar{y})^2}{Y'Y - n(\bar{y})^2}. \quad (13.1)$$

Недоліком коефіцієнта детермінації (13.1) є те, що він збільшується при введенні нових факторів, хоча це не завжди означає поліпшення регресійної моделі.

У цьому випадку краще використати *скоректований коефіцієнт детермінації*:

$$\hat{R}^2 = 1 - \frac{n-1}{n-p-1}(1-R^2). \quad (13.2)$$

За відомим коефіцієнтом детермінації визначимо *значущість рівняння регресії*. Обчислимо статистику:

$$F = \frac{R^2(n-p-1)}{(1-R^2)p}, \quad (13.3)$$

де p – число параметрів.

За таблицею Фішера–Снедекора знаходимо $F_{\alpha; p; n-p-1}$, якщо $F > F_{\alpha; p; n-p-1}$, то рівняння регресії значуще.

13.3. Мультиколінеарність

Мультиколінеарністю називається висока залежність (корельованість) різних факторів.

Якщо між факторами є тісний зв'язок, то визначник матриці $(X'X)$ дуже малий. Це призводить до великих помилок при оцінці параметрів рівняння регресії. В остаточному підсумку виходить незначуще рівняння регресії, що не має реального змісту.

Усунути або зменшити *мультиколінеарність* можна **методом покрокового відбору найбільш інформативних факторів**. Цю процедуру розглянемо на прикладі.

Приклад 13.1 За даними $n = 20$ сільськогосподарських районів урожайність Y (ц/га) залежить від факторів: X_1 – число тракторів на 100 га; X_2 – число комбайнів на 100 га; X_3 – число знарядь обробки ґрунту на 100 га; X_4 – число добрив на 1 га (т/га); X_5 – кількість засобів хімічного захисту на 1 га (ц/га).

Таблиця 13.1

i	y_i	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}
1	9,7	1,59	0,26	2,05	0,32	0,14
2	8,4	0,34	0,28	0,46	0,59	0,66
...
20	8,7	1,36	0,26	0,17	0,99	0,42

За даними табл. 13.1 складемо рівняння регресії:

$$\hat{y} = 3,515 - 0,006x_1 + 15,5x_2 + 0,11x_3 + 4,475x_4 - 2,932x_5.$$

(5,41) (0,6) (21,59) (0,85) (1,54) (3,09)

Примітка. У дужках наведені середньоквадратичні відхилення.

Визначимо значущість коефіцієнтів рівняння регресії:

$$t_{b_j} = \frac{|b_j|}{S_{b_j}} \cdot t_{b_0} = 0,65; t_{b_1} = 0,01; t_{b_2} = 0,72; t_{b_3} = 0,13; t_{b_4} = 2,91; t_{b_5} = 0,95.$$

За таблицею знаходимо: $t_{0,05; 20-5-1=14} = 2,14$.

Таким чином, одержуємо, що тільки $t_{b_4} = 2,91 > 2,14 = t_{0,05; 14}$. Тоді значущим є коефіцієнт b_4 .

За F -критерієм Фішера визначаємо значущість рівняння регресії.

$$F = \frac{R^2(n-p-1)}{(1-R^2)p},$$

де R^2 – коефіцієнт детермінації, що дорівнює 0,517; $n = 20$; p – кількість факторів, що дорівнює 5.

Виходить, що $F = 3$.

За табл. $F_{0,05; 14} = 2,96$. Оскільки $F = 3 > 2,96 = F_{0,05; 14}$, то рівняння регресії значуще, але воно визначає залежність тільки від x_4 .

Проведемо *покроковий відбір* інформативних факторів.

Складемо таблицю парних коефіцієнтів кореляції за формулою

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n \cdot S_x \cdot S_y}.$$

Таблиця 13.2

	Y	X_1	X_2	X_3	X_4	X_5
Y	1				0,58	
X_1		1	0,85	0,98		
X_2		0,85	1	0,88		0,46
X_3		0,98	0,88	1		
X_4	0,58				1	0,57
X_5			0,46		0,57	1

Примітка. У таблицю занесені тільки значущі коефіцієнти кореляції.

Мультиколінеарність має місце між факторами x_1, x_2, x_3 попарно.

Для усунення мультиколінеарності застосуємо *процедуру покрокового відбору* найбільш інформативних змінних.

Перший крок: найбільший коефіцієнт детермінації має місце для x_4

$$R_{y,4}^2 = (0,58)^2 = 0,336.$$

Скоректований коефіцієнт детермінації: $\tilde{R}_{y,4}^2 = 0,299$.

Другий крок: обчислюємо коефіцієнт детермінації для різних пар факторів $(X_4 X_j)$ ($j = 1, 2, 3, 5$) за формулою

$$R^2 = \frac{b'X'Y' - n(\bar{y})^2}{X'Y - n(\bar{y})^2}. \quad (13.4)$$

Найбільшим виявився $R_{Y \cdot X_4 \cdot X_3}^2 = 0,483$.

$$\hat{R}_{Y \cdot X_4 \cdot X_3}^2 = 0,422.$$

Третій крок: за формулою (13.4) обчислюємо коефіцієнт детермінації для різних трійок факторів (X_4, X_3, X_j) ($j = 1, 2, 5$).

Максимальним коефіцієнтом детермінації виявився

$$R_{Y \cdot X_4, X_3, X_5}^2 = 0,513.$$

$$\hat{R}_{Y \cdot X_4, X_3, X_5}^2 = 0,422.$$

Висновок: оскільки скоректований коефіцієнт детермінації не збільшився на третьому кроці, то в регресійній моделі досить обмежитися лише двома раніше відібраними факторами (X_4, X_3) .

Знову розраховане рівняння регресії має вигляд

$$\hat{y} = 7,29 + 3,48x_3 + 3,48x_4.$$

(0,66) (0,13) (1,07)

Висновок: усі коефіцієнти регресії (b_0, b_3, b_4) значущі, отже, рівняння регресії теж значуще.

Одержувані за допомогою покрокової процедури набори факторів є найбільш оптимальними або близькими до них.

13.4. Лінійні регресійні моделі зі змінною структурою. Фіктивні змінні

На практиці виникає необхідність досліджувати вплив якісних факторів, що мають два або кілька рівнів (градацій). Якісні ознаки можуть істотно впливати на структуру лінійних зв'язків між змінними й приводити до стрибкоподібної зміни параметрів регресійної моделі. У цьому випадку говорять про дослідження регресійної моделі зі змінною структурою.

Для опису таких змінних уводять **фіктивні змінні**, тобто використовують *бінарні величини* (0 та 1).

Наприклад, стать:

$$z = \begin{cases} 1, & \text{якщо Ч;} \\ 0, & \text{якщо Ж.} \end{cases}$$

рівень освіти: середня школа, технікум, вища освіта:

$$z = \begin{cases} 1, & \text{якщо в/о;} \\ 0 & \end{cases};$$

$$z = \begin{cases} 1, & \text{якщо Т;} \\ 0. & \end{cases}$$

Приклад 13.2. Необхідно дослідити залежність між результатами вступних і курсових екзаменів з математики. X – число розв’язаних задач на вступних іспитах (завдання – 10 задач); Y – число розв’язаних задач на курсових іспитах (завдання – 7 задач); n – число студентів, $n = 12$.

Таблиця 13.3

i	x_i	y_i	Стать	i	x_i	y_i	Стать
1	10	6	Ч	7	6	3	Ж
2	6	4	Ж	8	7	4	Ч
3	8	4	Ч	9	9	7	Ч
4	8	5	Ж	10	6	3	Ж
5	6	4	Ж	11	5	2	Ч
6	7	7	Ч	12	7	3	Ж

Рівняння регресійної моделі має вигляд

$$y_i = \beta_0 + \beta_1 x_i + \alpha_1 z_{i_1} + \varepsilon_i.$$

За вихідним даними отримано оцінку цієї моделі:

$$\hat{y} = -1,165 + 0,743x + 0,466z_1,$$

$$\text{де } z_1 = \begin{cases} 1, & \text{якщо Ч;} \\ 0, & \text{якщо Ж.} \end{cases}$$

За F -критерієм на рівні $\alpha = 0,05$ це рівняння регресії значуще.

Висновок: на курсових іспитах студент розв'язує на 0,5 задачі більше, ніж студентка.

За t -критерієм коефіцієнт регресії α_1 при фіктивній змінній z_1 виявився незначущим. Це можливо, якщо:

- 1) недостатньо вихідних даних;
- 2) гіпотеза про те, що стать впливає на кількість розв'язаних задач, неправильна;
- 3) абітурієнт міг мати освіту (технікум, ПТУ, середня школа).

Тоді регресійна модель має вигляд

$$y_i = \beta_0 + \beta_1 x_1 + \alpha_1 z_{i1} + \alpha_{21} z_{i21} + \alpha_{22} z_{i22} + \varepsilon_i,$$

$$z_{21} = \begin{cases} 1, & \text{якщо абітурієнт закінчив с/ш;} \\ 0, & \end{cases}$$

$$z_{22} = \begin{cases} 1, & \text{якщо абітурієнт закінчив с/ш;} \\ 0. & \end{cases}$$

Якщо є дві вибірки, одна з яких невелика, та виникає завдання про їхнє об'єднання, то використовується **критерій Г. Чоу**.

Алгоритм критерію Г. Чоу:

- 1) За кожною із вибірок будуються лінійні регресійні моделі.

$$y_i = \beta'_0 + \sum_{j=1}^p \beta'_j \cdot x_{ij} + \varepsilon'_i, \quad i = 1, 2, \dots, n_1$$

$$y_i = \beta''_0 + \sum_{j=1}^p \beta''_j \cdot x_{ij} + \varepsilon''_i, \quad i = n + 1, \dots, n_1 + n_2.$$

- 2) Висувається гіпотеза:

$$\beta' = \beta'', \quad D(\varepsilon') = D(\varepsilon'').$$

- 3) Обчислюється статистика:

$$F = \frac{\left(\sum_{i=1}^n e_i^2 - \sum_{i=1}^{n_1} e_i^2 - \sum_{i=n_1+1}^n e_i^2 \right) (n - 2p - 1)}{\left(\sum_{i=1}^{n_1} e_i^2 + \sum_{i=n_1+1}^n e_i^2 \right) (p + 1)}, \quad (13.5)$$

де $\sum_{i=1}^n e_i^2, \sum_{i=1}^{n_1} e_i^2, \sum_{i=n_1+1}^n e_i^2$ – залишкові суми квадратів відхилень об'єднаної, першої й другої вибірок, $n = n_1 + n_2$.

4) За таблицею F -критерію знаходимо $F_{\alpha; p+1; n-2p-2}$, де p – число змінних; n – число спостережень.

5) Якщо $F > F_{\alpha; p+1; n-2p-2}$, то висунута гіпотеза відкидається.

Приклад 13.3. За даними прикладу 13.2 маємо дві вибірки:

для юнаків: $n_1 = 6; (10;6), (8;4), (7;7), (7;4), (9;7), (5;2);$

для дівчат: $n_2 = 6; (6;4), (8;5), (6;4), (6;3), (6;3), (7;3).$

За формулою (13.5) $F = 0,21$.

За таблицею знаходимо: $F_{0,05; 2; 8} = 4,46$.

Оскільки $F = 0,21 < 4,46$, то висунута гіпотеза приймається. Це означає, що вплив статі студента на результати іспитів несуттєвий.

Контрольні запитання

1. Довірчі інтервали для коефіцієнтів та функції множинної регресії.
2. Оцінка значущості рівняння множинної регресії.
3. Мультиколінеарність змінних, методи її усунення або зменшення.
4. Лінійні моделі зі змінною структурою. Фіктивні змінні.
5. Критерій Г. Чоу.

14. НЕЛІНІЙНІ РЕГРЕСІЙНІ МОДЕЛІ. ЛІНЕАРИЗАЦІЯ МОДЕЛІ. КОЕФІЦІЄНТИ ЧАСТКОВОЇ ЕЛАСТИЧНОСТІ. ЧАСТКОВА КОРЕЛЯЦІЯ

14.1. Нелінійні регресійні моделі. Лінеаризація моделі. Коефіцієнти часткової еластичності

Нехай модель регресії має вигляд

$$y_i = \beta_0 + \beta_1 x_{i1}^2 + \beta_2 \sqrt{x_{i2}} + \varepsilon_i.$$

Це рівняння є *нелінійним за змінними*.

Уведемо нові змінні:

$$z_1 = x_1^2, z_2 = \sqrt{x_2}.$$

Тоді $y_i = \beta_0 + \beta_1 z_{i1} + \beta_2 z_{i2} + \varepsilon_i$.

Недоліком цієї моделі є труднощі, що виникають при інтерпретації кінцевих результатів.

Розглянемо регресійні моделі, які *нелінійні за параметрами*.

Мультиплікативна (степенева) модель:

$$y_i = \beta_0 \cdot x_{i1}^{\beta_1} \cdot x_{i2}^{\beta_2} \cdot \varepsilon_i.$$

Експонентна модель:

$$y_i = e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i}.$$

Логарифмуючи обидві частини цих рівнянь, одержимо лінійну модель щодо логарифмів змінних.

$$\ln y_i = \ln \beta_0 + \beta_1 \ln x_{i1} + \beta_2 \ln x_{i2} + \ln \varepsilon_i.$$

Для цього рівняння можна застосувати метод найменших квадратів (МНК), але необхідно, щоб вектор збурень ε мав *логарифмічно нормальний розподіл*.

Задача залежності обсягу виробництва від витрат капіталу й праці називається *задачею Кобба–Дугласа*.

$$Y = AK^\alpha L^\beta \varepsilon, \quad (14.1)$$

де Y – обсяг виробництва; A – коефіцієнт пропорційності; K – витрати капіталу; L – витрати праці; ε – вектор збурювань; α, β – **коефіцієнти часткової еластичності** обсягу виробництва за витратами капіталу й праці відповідно.

Коефіцієнтом часткової еластичності $E_{x_i}(y)$ функції $y = f(x_1, x_2, \dots, x_n)$ називається межа відношення відносного часткового збільшення функції до відносного збільшення цієї змінної при $\Delta x_i \rightarrow 0$, тобто

$$E_x = \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta xy}{y} : \frac{\Delta x}{x} \right) = \frac{x}{y} \cdot y'_x,$$

де y'_x – коефіцієнт еластичності; $E_K = \alpha$; $E_L = \beta$.

Лінеаризація моделі (8.1) проходить шляхом логарифмування:

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L + \ln \varepsilon.$$

При розширенні виробництва (збільшення витрат капіталу й праці в декілька разів приводить до збільшення обсягу виробництва на таке ж число разів) $\alpha + \beta = 1$.

У цьому випадку $Y = AK^\alpha L^{1-\alpha} \varepsilon$:

$$\frac{Y}{L} = A \left(\frac{K}{L} \right)^\alpha \varepsilon - \text{модель розширення виробництва},$$

де $\frac{Y}{L}$ – *продуктивність праці*; $\frac{K}{L}$ – *капіталоозброєність виробництва*.

Лінеаризація цієї моделі здійснюється шляхом логарифмування.

Функція Кобба–Дугласа з урахуванням технічного прогресу має вигляд

$$Y = AK^\alpha L^\beta e^{\theta t} \varepsilon,$$

де θ – параметр темпу приросту обсягу виробництва завдяки технічному прогресу; t – час.

14.2. Часткова кореляція

Частковим коефіцієнтом кореляції між змінними X_i, X_j називається вираз

$$r_{ij.12\dots p} = \frac{-q_{ij}}{\sqrt{q_{ii}q_{jj}}},$$

де q_{ij} – алгебричне доповнення елементів r_{ij} матриці парних коефіцієнтів кореляції.

$$q_p = \begin{pmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{pmatrix},$$

$$\text{де } r_{ij} = \frac{\sum_{i=1}^n (x_i - \bar{x}_i)(x_j - \bar{x}_j)}{n \cdot S_{x_i} \cdot S_{x_j}}.$$

Для трьох змінних

$$r_{ij.k} = \frac{r_{ij} - r_{ik}r_{jk}}{\sqrt{(1 - r_{ik}^2)(1 - r_{jk}^2)}}.$$

Значущість часткового коефіцієнта кореляції оцінюється так само, як і для парного коефіцієнта кореляції, припускаючи при цьому $n' = n - p + 2$.

Контрольні запитання

1. Види нелінійних регресійних моделей.
2. Функція Кобба – Дугласа та її види.
3. Коефіцієнти часткової еластичності та кореляції.

15. УЗАГАЛЬНЕНІ ЕКОНОМЕТРИЧНІ МОДЕЛІ

15.1. Узагальнена лінійна модель множинної регресії

Узагальнена лінійна модель множинної регресії (*Generalized Linear Multiple Regression Model*)

$$Y = X\beta + \varepsilon, \quad (15.1)$$

описується наступною системою співвідношень і умов:

1. ε – випадковий вектор; X – не випадкова (детермінована) матриця;
2. $M(\varepsilon) = 0_n$;
3. $\Sigma_\varepsilon = M(\varepsilon\varepsilon') = \Omega$, де Ω – позитивно визначена матриця;
4. $r(X) = p + 1 < n$,

де p – число пояснюючих змінних; n – число спостережень.

Порівнюючи узагальнену модель із класичною, бачимо, що вона відрізняється від класичної тільки видом коваріаційної матриці: замість $\Sigma_\varepsilon = \sigma^2 E_n$ для класичної моделі маємо $\Sigma_\varepsilon = \Omega$ для узагальненої. Це означає, що на відміну від класичної в узагальненій моделі коваріації та дисперсії пояснюючих змінних можуть бути довільними. У цьому полягає суть узагальнення регресійної моделі.

Для оцінки параметрів моделі (15.1) можна застосувати звичайний метод найменших квадратів.

Оцінка $b = (X'X)^{-1} X'Y$ залишається правдивою й у випадку узагальненої моделі. Оцінка b як і раніше незміщена й спроможна.

Однак отримана раніше формула для коваріаційної матриці вектора оцінок Σ_b виявляється неприйнятною в умовах узагальненої моделі.

Дійсно, одержимо для узагальненої моделі

$$\Sigma_b = (X'X)^{-1} X' M(\varepsilon\varepsilon') X (X'X)^{-1} = (X'X)^{-1} X' \Omega X (X'X)^{-1}, \quad (15.2)$$

у той час як для класичної моделі

$$\Sigma_b = \sigma^2 (X'X)^{-1}. \quad (15.3)$$

Знайдемо математичне очікування залишкової суми квадратів

$$\sum_{i=1}^n e_i^2 = e'e.$$

Використовуючи перетворення, можна показати, що для узагальненої моделі

$$M(e'e) = \text{tr}[(E_n - X(X'X)^{-1}X')\Omega], \quad (15.4)$$

$$\text{тобто} \quad M(s^2) = M\left(\frac{ee'}{n-p-1}\right) = \frac{\text{tr}[(E_n - X(X'X)^{-1}X')\Omega]}{n-p-1}, \quad (15.5)$$

де символ tr означає слід відповідної матриці.

Отже, якщо як оцінку коваріаційної матриці Σ_b в співвідношенні (15.3) замінити σ^2 на s^2 , тобто взяти матрицю $\Sigma_b = s^2(X'X)^{-1}$, то її математичне очікування

$$M(\Sigma_b) = M(s^2)(X'X)^{-1} = \frac{\text{tr}[(E_n - X(X'X)^{-1}X')\Omega]}{n-p-1}(X'X)^{-1} \quad (15.6)$$

у загальному випадку не збігається з коваріаційною матрицею, визначеною співвідношенням (15.2). Це означає, що звичайний метод найменших квадратів в узагальненій лінійній регресійній моделі дає зміщену оцінку коваріаційної матриці Σ_b вектора оцінок b .

Оцінка b хоча й буде спроможною, але не буде оптимальною за теоремою Гаусса–Маркова. Для одержання найбільш ефективної оцінки потрібно використовувати іншу оцінку, одержувану так званим узагальненим методом найменших квадратів.

15.2. Узагальнений метод найменших квадратів

Економетрична модель, в якій спостерігається явище гетероскедастичності, є узагальненою моделлю і для оцінювання її параметрів слід використовувати не метод найменших квадратів (МНК), а узагальнений метод найменших квадратів (метод Ейткена, або скорочено УМНК).

Ідея УМНК полягає у знаходженні оцінок матриці параметрів b моделі з використанням додатково визначеної діагональної матриці Ω , за допомогою якої коригується вхідна інформація.

Матриця Ω має вигляд:

$$\Omega = \begin{pmatrix} 1/\lambda_1 & 0 & 0 & \dots & 0 \\ 0 & 1/\lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 1/\lambda_n \end{pmatrix}, \quad (15.7)$$

де λ_i – параметри, які обчислюються з використанням гіпотез:

а) дисперсія залишків σ_ε^2 пропорційна до змін пояснювальної змінної x_i , тоді $\lambda_i = \frac{1}{x_i} (i = \overline{1, n}, j = \overline{1, m})$;

б) дисперсія залишків пропорційна до зміни квадрата пояснювальної змінної x_i^2 , тоді $\lambda_i = \frac{1}{x_i^2}$;

в) дисперсія залишків пропорційна до зміни квадрата залишків за модулем $|\varepsilon_i|^2$, тоді $\lambda_i = \{|\varepsilon_i|\}^2$.

Оскільки матриця $\Omega = P'P$, то матриця P має вигляд:

$$P = \begin{pmatrix} 1/\sqrt{\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & 1/\sqrt{\lambda_2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 1/\sqrt{\lambda_n} \end{pmatrix}; \quad P^{-1} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \sqrt{\lambda_n} \end{pmatrix}.$$

За наявності гетероскедастичності узагальнена модель має вигляд

$$Y^* = X^* \beta + \varepsilon^*, \quad (15.8)$$

де $Y^* = P^{-1}Y$; $X^* = P^{-1}X$; $\varepsilon^* = P^{-1}\varepsilon$.

Використання для узагальненої моделі (15.8) МНК приводить до такого оператора оцінювання параметрів УМНК:

$$b = (X^{*'} X^*)^{-1} X^{*'} Y^* = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y. \quad (15.9)$$

15.2. Поняття гомо- та гетероскедастичності

Однією з чотирьох необхідних умов для застосування МНК при оцінюванні параметрів економетричної моделі є вимога постійної дисперсії залишків для кожного спостереження, тобто $M(\varepsilon \varepsilon') = \sigma_\varepsilon^2 E$. Ця властивість незмінної дисперсії в спостереженнях називається *гомоскедастичністю*.

Якщо дисперсія залишків змінюється для кожного спостереження або групи спостережень, тобто $M(\varepsilon \varepsilon') = \sigma_\varepsilon^2 \Omega$, то це явище називається *гетероскедастичністю*. Тут позначено: σ_ε^2 – дисперсія залишків, яка є невідомим параметром; Ω – відома симетрична додатно визначена матриця. Зауважи-

мо, що обидва терміни – гомоскедастичність та гетероскедастичність – запропоновані відомим російським вченим-статистиком А.А. Чупровим.

За наявності гетероскедастичності залишків ε оцінки параметрів моделі МНК будуть незміщеними, обґрунтованими, але неефективними. При цьому формулу для обчислення стандартної помилки спостережень застосовувати не можна.

За наявності гетероскедастичності в простій економетричній моделі, тобто $Y = b_0 + b_1X + \varepsilon$, щоб оцінити параметри МНК, достатньо змінити специфікацію моделі. Коли будується модель множинної регресії з багатьма змінними, то таке перетворення значно ускладнюється. Тому спочатку треба одним із методів визначити наявність гетероскедастичності, а потім оцінити параметри моделі із застосуванням спеціального підходу, який буде розглянутий пізніше.

15.3. Методи визначення гетероскедастичності. Критерій μ

Цей метод використовується при великій кількості сукупності спостережень n , яких може бути удвічі більше, ніж оцінюваних параметрів. Вхідні дані залежної змінної Y розбиваються на k груп з номерами $r=1\dots k$. За кожною групою спостережень розраховується сума квадратів відхилень:

$$S_r = \sum_{i=1}^k (y_{ir} - \bar{y}_r)^2.$$

Обчислюється сума квадратів відхилень у цілому для сукупності спостережень:

$$\sum_{r=1}^k S_r = \sum_{i=1}^{n_r} \sum_{r=1}^k (y_{ir} - \bar{y}_r)^2,$$

де n_r – кількість спостережень у кожній групі r . Розраховується параметр λ :

$$\lambda = \prod_{i=1}^k \left(\frac{S_r}{n_r} \right)^{n_r/2} / \left(\frac{\sum_{r=1}^k S_r}{n} \right)^{n/2}$$

де Π – добуток k виразів за цим позначенням.

Обчислюється критерій

$$\mu = -2 \ln \lambda,$$

який наближено відповідає розподілу χ^2 при ступені волі $(k-1)$ та рівні довіри $(1-\alpha)$. Якщо $\mu > \chi^2_{\text{табл}}$, то у множині спостережень, що розглядається, має місце гетероскедастичність.

Параметричний тест Гольдфелда–Квандта. Метод застосовується при досить великій кількості спостережень n , коли дисперсія залишків зростає пропорційно квадрату однієї з незалежних змінних моделі. Тест починається з упорядкування за зменшенням незалежних змінних x_j ($j=1...m$) відносно тієї змінної, яка підозрюється на гетероскедастичність. Виключають c середніх (у цьому упорядкуванні) спостережень. Оптимальне значення c за дослідженням авторів тесту становить

$$c = \frac{4}{15}n.$$

Будуються дві економетричні моделі на основі МНК з розміром спостережень $(n - c)/2$ для першої моделі і таким же розміром для другої. Обчислюється сума квадратів залишків за першою S_1 та другою S_2 моделями:

$$S_1 = \varepsilon'_1 \varepsilon_1; S_2 = \varepsilon'_2 \varepsilon_2.$$

де ε_1 та ε_2 - залишки моделей (1) та (2).

Обчислюється критерій

$$R^* = \frac{S_2}{S_1},$$

який наближено відповідає F -критерію Фішера. За статистичними таблицями знаходиться значення $F_{\text{табл}}$ для ступенів волі $(n-c-2m_1)/2$, $(n-c-2m_1)/2$ і вибраного рівня довіри, де m_1 – загальна кількість оцінюваних параметрів у моделі. Якщо $R^* > F_{\text{табл}}$, то це свідчить про наявність гетероскедастичності.

Непараметричний тест Гольдфелда–Квандта базується на розгляді графічно зображеної залежності залишків ε_i від упорядкованих спостережень за значеннями x_j ($j=1...m$). Якщо для всіх значень змінної x_j залишки розподіляються нерівномірно і без певної закономірності, то дисперсія залишків є змінною величиною і спостерігається явище гетероскедастичності.

Тест Глейсера відповідає знаходженню параметрів регресійної моделі за допомогою МНК при використанні абсолютних значень залишків у

функції від незалежної змінної x_j , яка може викликати зміну дисперсії σ_ε^2 .
Для цього використовується один з таких видів функцій:

$$|\varepsilon| = b_0 + b_1 x_j; |\varepsilon| = b_0 + b_1 x_j^{-1}; |\varepsilon| = b_0 + b_1 x_j^{\frac{1}{2}}.$$

Рішення про гетероскедастичність на підставі значень параметрів b_0 і b_1 таке: чистій гетероскедастичності відповідає значення параметрів $b_0 = 0, b_1 \neq 0$, змішаній гетероскедастичності – $b_0 \neq 0, b_1 = 0$.

Контрольні запитання

1. Узагальнена лінійна модель множинної регресії.
2. Алгоритм узагальненого методу найменших квадратів.
3. Поняття гомо- та гетероскедастичності. Критерій μ .
4. Параметричний та непараметричний тести Гольдфелда–Квандта.

Тест Глейсера.

16. ЧАСОВІ РЯДИ

16.1. Основні поняття

Часові ряди – це сукупність значень будь-якої ознаки за кілька послідовних моментів або періодів часу.

Окремі значення цієї ознаки називаються *рівнями ряду*. Рівні часового ряду визначаються безліччю факторів, які можна розбити на кілька груп:

1. Фактори, що формують тенденцію ряду або його **тренд** (T).
2. Циклічні коливання (S).
3. Випадкові фактори (E).

Найбільш часто використовують:

- 1) **адитивну модель**

$$Y = T + S + E;$$

- 2) **мультиплікативну модель**

$$Y = T \cdot S \cdot E.$$

16.2. Автокореляція рівнів часового ряду

У загальному випадку наступні рівні часового ряду залежать від попередніх. Кореляційна залежність між послідовними рівнями ряду називається *автокореляцією рівнів ряду*. Кількісно вона оцінюється за допомогою лінійного парного коефіцієнта кореляції між рівнями, зрушеними один відносно одного. Крок зрушення називається *лагом*.

Приклад 16.1. Є дані про середні витрати на кінцеве споживання за вісім років (тис. у.о.) (табл. 16.1).

Таблиця 16.1

t	1	2	3	4	5	6	7	8
y_t	7	8	8	10	11	12	14	16
y_{t-1}	–	7	8	8	10	11	12	14

Коефіцієнт кореляції обчислюється за формулою

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}.$$

Тоді ця формула набуває вигляду

$$r_1 = \frac{\sum_{t=2}^n (y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^n (y_t - \bar{y}_1)^2 \sum_{t=2}^n (y_{t-1} - \bar{y}_2)^2}} \quad \bar{y}_1 = \sum_{t=2}^n y_t / (n-1);$$

$$\bar{y}_2 = \sum_{t=2}^n y_{t-1} / (n-1).$$

Величина r_1 називається коефіцієнтом автокореляції рівнів ряду першого порядку (лаг = 1):

$$\bar{y}_1 = \frac{8+8+\dots+16}{7} = 11,29;$$

$$\bar{y}_2 = \frac{7+8+\dots+14}{7} = 10;$$

$$r_1 = 0,975.$$

Для коефіцієнта автокореляції другого порядку використовуємо такі зрушені ряди (табл. 16.2):

Таблиця 16.2

t	1	2	3	4	5	6	7	8
y_t	7	8	8	10	11	12	14	16
y_{t-2}	—	—	7	8	8	10	11	12

$$r_2 = \frac{\sum_{t=3}^n (y_t - \bar{y}_3)(y_{t-2} - \bar{y}_4)}{\sqrt{\sum_{t=3}^n (y_t - \bar{y}_3)^2 \sum_{t=3}^n (y_{t-2} - \bar{y}_4)^2}} \quad \bar{y}_3 = \frac{\sum_{t=3}^n y_t}{n-2};$$

$$\bar{y}_4 = \frac{\sum_{t=3}^n y_{t-2}}{n-2}.$$

Обчислений $r_2 = 0,973$.

Аналогічно обчислюємо коефіцієнти автокореляції 3-го, 4-го та інших порядків.

Для одержання статистично достовірних результатів вважають, що максимальний лаг не повинен перевищувати величину $(n : 4)$.

Послідовність коефіцієнтів автокореляції називається *автокореляційною функцією*, а її графік називається *корелограмою*.

Аналіз автокореляційної функції дозволяє зробити деякі висновки:

1. Якщо найбільшим виявився r_1 , то часовий ряд містить тільки тенденцію.
2. Якщо найбільшим виявився r_τ , то ряд містить циклічне коливання з періодом τ .
3. Якщо жоден із коефіцієнтів автокореляції не є значущим, то ряд:
 - або містить тільки випадкову складову;
 - або містить сильну нелінійну тенденцію.

16.3. Моделювання часового ряду

Моделювання часового ряду зводиться до розрахунку значень T , S , E для кожного рівня ряду.

Цей процес включає наступні кроки:

- 1) вирівнювання часового ряду методом ковзної середньої (табл.16.3).

Таблиця 16.3

t	1	2	3	4	5	6	7	8
y_t	213	171	291	309	317	362	351	361
\tilde{y}_t	–	225	241	305,7	329,3	336,3	358	–

$$\tilde{y}_2 = \frac{y_1 + y_2 + y_3}{3} = (213 + 171 + 291)/3 = 225; \quad \tilde{y}_3 = \frac{y_2 + y_3 + y_4}{3} = 241;$$

- 2) розрахунок значень циклічного компонента S ;
- 3) усунення S з вихідних рівнів часового ряду;
- 4) аналітичне вирівнювання рівнів ряду (одержання рівняння тренда) і розрахунок значень T з використанням отриманого рівняння тренда;
- 5) розрахунок отриманих за моделлю значень $T + S$ або $T \cdot S$;
- 6) розрахунок помилок.

Приклад 16.2. Є дані про споживання електроенергії в районі за чотири роки (млн.кВт/г) (табл. 16.4).

Таблиця 16.4

Номер кварталу t	Споживання електроенергії, y_t	Σ за чотири квартали	Ковзне се- реднє за чо- тири квар- тали	Центроване ковзне середнє	Циклічний компонент S_t
1	2	3	4	5	6
1	6	—	—	—	—
2	4,4	24,4	6,1	—	—
3	5	25,6	6,4	6,25	-1,25
4	9	26	6,5	6,45	2,55
5	7,2	27	6,75	6,625	0,575
6	4,8	28	7	6,875	-2,075
...
16	10,8	—	—	—	—

З аналізу вихідних рівнів ряду видно, що амплітуда коливань приблизно постійна, тому необхідно скористатися *адитивною моделлю*:

$$Y = T + S + E.$$

КРОК 1: Вирівнюємо вихідні рівні методом ковзної середньої. Для цього:

а) підсумуємо рівні ряду послідовно за кожні чотири квартали зі зрушенням на 1 квартал (графа 3.4: 3 – номер стовпця, 4 – номер рядка);

б) розділивши отримані суми на 4, знайдемо значення ковзної середньої (графа 4.4). Ці значення вже не містять циклічного компонента;

в) приведемо ці значення у відповідність до фактичних моментів часу, для чого знайдемо середнє значення із двох сусідніх значень – центровані значення ковзної середньої (графа 5.4).

КРОК 2: Знайдемо оцінки циклічного компонента S_t як різниці між y_t і центрованими значеннями ковзної середньої (графа 6.4). У моделях із циклічним компонентом звичайно передбачається, що циклічні впливи за період взаємопогашаються. В адитивній моделі це означає, що сума значень \bar{S}_i за всіма кварталами повинна дорівнювати нулю.

$$\sum \bar{S}_i = 0,6 - 1,958 - 1,275 + 2,708 = 0,075.$$

Коригувальне виправлення визначається за формулою $K = \frac{0,075}{4} = 0,01875$.

Скоректовані значення S_i : $S_i = \bar{S}_i - K$. Отримані значення S_i подані в табл. 16.5

Таблиця 16.5

Показник	Рік	Номер кварталу (<i>i</i>)			
		I	II	III	IV
	1	–	–	–1,25	2,55
	2	0,575	–2,075	–1,1	2,7
	3	0,55	–2,025	–1,475	2,875
	4	0,675	–1,775	–	–
Разом за <i>i</i> -й квартал (за чотири роки)		1,8	–5,875	–3,825	8,125
Середня оцінка \bar{S}_i		0,6	–1,958	–1,275	2,708
Скоректована оцінка S_i		0,581	–1,977	–1,294	2,69

КРОК 3: Елімінуємо (видалимо) вплив циклічного компонента, віднімаючи його значення з рівнів часового ряду y_t :

$$(T + E)_i = y_i - S_i.$$

Таблиця 16.6

t	y_t	S	$T + E$	T	$T + S$	E	E^2
1	2	3	4	5	6	7	8
1	6	0,581	5,419	5,902	6,483	–0,483	0,2333
2	4,4	–1,977	6,337	6,088	4,111	0,289	0,0835
3	5	–1,294	6,294	6,275	4,981	0,019	0,0004
4	9	2,69	6,31	6,461	9,151	–0,151	0,0228
5	7,2	0,581	6,619	6,648	7,229	–0,29	0,0008
6	4,8	–1,977	6,777	6,834	4,857	–0,057	0,0032
...
16	10,8	2,69	8,11	8,698	11,388	–0,588	0,3457

КРОК 4: Визначимо компонент T . Для цього проведемо аналітичне вирівнювання часового ряду $T + E$ (графа 4.6: 4 – номер стовпця, 6 – номер

рядка) за допомогою лінійного тренду. За формулами розділу 10 для парної лінійної регресії:

$$T = 5,715 + 0,186t.$$

У цій формулі замість змінної x був підставлений час t . Із цього рівняння знайдемо значення T для кожного моменту часу (графа 5.6).

КРОК 5: Обчислимо значення абсолютних відхилень (графа 7.6) та їхні квадрати (графа 8.6). За даними графі 8.6 знайдемо коефіцієнт детермінації:

$$Q = \sum_t (y_t - \bar{y}_t)^2;$$

$$Q_e = \sum_t E^2 = 1,1.$$

Тоді $R^2 = \frac{1-1,1}{71,59} = 0,985$. Це означає, що адитивна модель пояснює 98,5 % загальної варіації рівнів часового ряду споживання електроенергії.

Приклад 16.3 Знайти прогноз на споживання електроенергії в першому півріччі наступного року.

$$T_{17} = 5,715 + 0,186 \cdot 17 = 8,877; T_{18} = 5,715 + 0,186 \cdot 18 = 9,063.$$

$$S_{17} = S_I = 0,581; S_{18} = S_{II} = -1,977.$$

Тоді $F = T_{17} + T_{18} + S_{17} + S_{18} = 16,544$ млн кВт/г.

Аналогічним чином будують *мультиплікативну модель*.

Контрольні запитання

1. Часовий ряд, його складові та моделі.
2. Автокореляція рівнів часового ряду, автокореляційна функція, її властивості.
3. Алгоритм моделювання часового ряду.

СПИСОК ЛІТЕРАТУРИ

1. Боровік О. В. Дослідження операцій в економіці : навч. посіб. / О. В. Боровік, Л. В. Боровік. – К. : Центр навчальної літератури, 2007. – 424 с.
2. Исследование операций в экономике : учеб. пособие для вузов / Н. Ш. Кремер, Б. А. Путко, И. М. Тришин, М. Н. Фридман. – М. : ЮНИТИ, 2006. – 407 с.
3. Вентцель Е. С. Исследование операций. Задачи, принципы, методология : учеб. пособие для студ. вузов / Е. С. Вентцель. – М. : Высш. шк., 2001. – 208 с.
4. Кутковецький В. Я. Дослідження операцій : навч. посібник / В. Я. Кутковецький. – К. : „Видавничий дім „Професіонал”, 2004. – 350 с.
5. Білоцерківський О. Б. Економетрія : навч.-метод. посібник / О. Б. Білоцерківський, Н. В. Ширяєва. – Х. : НТУ "ХП", 2008. – 80 с.
6. Клебанова Т. С. Эконометрия : учебное пособие / Т. С. Клебанова, Н. А. Дубровина, Е. В. Раевнева. – Х. : ИНЖЭК, 2005. – 160 с.
7. Кремер Н. Ш. Эконометрика : учебник для вузов / Н. Ш. Кремер, Б. А. Путко; под ред. проф. Н. Ш. Кремера. – М. : ЮНИТИ-ДАНА, 2005. – 311 с.
8. Доугерти К. Введение в эконометрику : учебник / К. Доугерти. – М. : ИНФРА-М, 2004. – 432 с.
9. Елисеева И. И. Эконометрика : учебник для вузов / И. И. Елисеева, С. В. Курышева, Т. В. Костеева; под ред. чл.-корр. РАН И. И. Елисеевой. – М. : Финансы и статистика, 2002. – 342 с.
10. Елисеева И. И. Практикум по эконометрике / И. И. Елисеева, С. В. Курышева, Н. М. Гордеенко; под ред. чл.-корр. РАН И. И. Елисеевой. – М. : Финансы и статистика, 2002. – 191 с.
11. Наконечний С. І. Економетрія : підручник / С. І. Наконечний, Т. О. Терещенко, Т. П. Романюк. – К. : КНЕУ, 2005. – 520 с.
12. Лугінін О. Є. Економетрія : навчальний посібник / О. Є. Лугінін, С. В. Білоусова, О. М. Білоусов. – К. : Центр навчальної літератури, 2005. – 252 с.

ЗМІСТ

Вступ.....	3
1 Концептуальні аспекти математичного моделювання економіки.....	4
2 Задачі лінійного програмування.....	11
3 Методи розв’язування задач лінійного програмування.....	23
4 Теорія двоїстості та аналіз лінійних моделей оптимізаційних задач...	32
5 Моделі цілочислового лінійного програмування.....	38
6 Моделі нелінійного програмування.....	44
7 Моделі динамічного програмування.....	53
8 Елементи теорії ігор.....	57
9 Предмет, методи та задачі економетрії. Особливості побудови економетричних моделей.....	65
10 Парний регресійний аналіз.....	69
11 Основні положення регресійного аналізу. Оцінка параметрів регресійної моделі. Теорема Гаусса–Маркова. Інтервальна оцінка функції регресії та її параметрів. Оцінка значущості рівняння регресії.....	72
12 Множинний регресійний аналіз.....	77
13 Довірчі інтервали для коефіцієнтів і функції регресії. Оцінка значущості рівняння множинної регресії. Мультиколінеарність. Лінійні регресійні моделі зі змінною структурою. Фіктивні змінні.....	83
14 Нелінійні регресійні моделі. Лінеаризація моделі. Коефіцієнти часткової еластичності. Часткова кореляція.....	91
15 Узагальнені економетричні моделі.....	94
16 Часові ряди.....	100
Список літератури.....	106

Навчальне видання

БЛОЦЕРКІВСЬКИЙ Олександр Борисович

ШИРЯЄВА Наталя Володимирівна

ЗАМУЛА Олексій Олександрович

ЕКОНОМІКО-МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Текст лекцій

для студентів спеціальностей

7.050206 "Менеджмент зовнішньоекономічної діяльності"

та 6.030508 "Фінанси"

Роботу до видання рекомендував проф. В. А. Міщенко

Редактор О. С. Самініна

План 2009 р. п. 114/

Підписано до друку 18.10.10. Формат $60 \times 84 \frac{1}{16}$. Папір офсет.

Друк - ризографія. Гарнітура Таймс. Ум. друк. арк. 5,7. Обл.-вид. арк. 6,4.

Наклад 100 прим. Зам № . Ціна договірна.

Видавничий центр НТУ "ХП".

Свідоцтво про державну реєстрацію ДК № 3657 від 24.12.2009р.

61002, Харків 2, вул. Фрунзе, 21

Друкарня НТУ "ХП". 62002, Харків 2, вул. Фрунзе, 21.